Title Page:

Analysis of gene network bifurcation during optic cup morphogenesis in zebrafish

Lorena Buono[1,2], Silvia Naranjo[1], Tania Moreno-Marmol[2], Berta de la Cerda[3], Rocío Polvillo[1], Francisco-Javier Díaz-Corrales[3], Ozren Bogdanovic[4,5], Paola Bovolenta[2,6#], Juan-Ramón Martínez-Morales[1#]

1 Centro Andaluz de Biología del Desarrollo- CABD (CSIC/UPO/JA), Seville 41013, Spain

2 Centro de Biología Molecular "Severo Ochoa," (CSIC/UAM)

3. Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER (CSIC/US/UPO/JA), Seville 41092, Spain.

4. Genomics and Epigenetics Division, Garvan Institute of Medical Research, Sydney 2010, New South Wales, Australia.

5. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, New South Wales, 2052, Australia

6 CIBERER, ISCIII, Madrid 28049, Spain

Correspondence: pbovolenta@cbm.csic.es, jrmarmor@upo.es,

Summary

Sight depends on the tight cooperation between photoreceptors and pigmented cells. Both derive from common progenitors in which a single gene regulatory network (GRN) bifurcates into the neural retina (NR) and retinal-pigmented epithelium (RPE) programs. Although genetic studies have identified upstream nodes controlling these networks, their regulatory logic remains poorly investigated. Here, we characterize transcriptome dynamics (RNA-seq) and chromatin accessibility (ATAC-seq) in segregating NR/RPE populations in zebrafish. Analysis of active cis-regulatory modules and enriched transcription factor (TF) motives suggest extensive network redundancy and context-dependent TF activity. Downstream targets identification highlights an early recruitment of desmosomal genes in the flattening RPE, revealing Tead factors as upstream regulators. Investigation of GRNs dynamics uncovers an unexpected sequence of TF recruitment during RPE specification, which is conserved in humans. This systematic interrogation of the NR/RPE bifurcation should improve both genetic counselling for eye disorders and hiPSCs-to-RPE differentiation protocols for cell-replacement therapies in degenerative diseases.

Keywords: Eye development, Gene regulatory network (GRN), cis-regulatory modules, Retinal degenerative disease, neural retina (NR), retinal-pigmented epithelium (RPE).

Running title: Retinal network bifurcation in zebrafish

Introduction

Darwin in *The Origin of Species* already hinted to the possibility of a single photoreceptor escorted by a pigmented cell, as the ancestral structure to all animal eyes. Several authors have further elaborated this notion during the last decades (Gehring, 2014; Vopalensky and Kozmik, 2009; Arendt and Wittbrodt, 2001). An intimate association between photoreceptors and pigmented ancillary cells is found in the eyes of all metazoan groups (Land and Nilsson, 2002), even in cnidarians (Kozmik, et al., 2008). Thus, the origin of this cellular tandem can be traced back to a common ancestor dating to at least 650 mya. Such remarkable evolutionary conservation can be explained by the fact that vision depends on the close physical and physiological interaction between the two cell types (Strauss, 2005). Photoreceptors, with a very active metabolism and a constant exposure to light, are particularly prone to degeneration due to DNA damage and oxidative stress. Pigmented cells protect photoreceptors by maintaining their homeostasis through growth factors secretion, visual pigments recycling and outer-segments phagocytosis (Letelier, et al., 2017).

Fate map experiments in vertebrates have shown that the precursors of the neural retina (NR) and retinal pigmented epithelium (RPE) derive from the undifferentiated eye field (Li, et al., 2000). In the permanently growing retinas of teleost fish, stem cells located at the ciliary marginal zone are capable of producing both cell types (Tang, et al., 2017). Furthermore, in tetrapods retinal precursors retain certain potential to trans-differentiate from the pigmented to the neural retina identity (Pittack, et al., 1991; Coulombre and Coulombre, 1965), and *vice versa* (Rowan, et al., 2004).

In zebrafish, the RPE and NR presumptive domains start differentiating at the optic vesicle stage from the medial (ML) and lateral (LL) epithelial layers, respectively (Kwan, et al., 2012; Li, et al., 2000). The specification of both retinal domains occurs simultaneously to the folding of the vesicle into a bi-layered cup, and entails profound cell shape changes in each domain. Precursors at the LL elongate along their apico-basal axis, differentiate as NR, and constrict basally to direct the folding of the retinal neuroepithelium (Nicolas-Perez, et al., 2016; Ivanovitch, et al., 2013). In contrast, precursors at the ML either acquire a squamous epithelial shape and differentiate as RPE, or flow into the LL to contribute to the NR domain (Moreno-Marmol, et al., 2018; Sidhaye and Norden, 2017; Heermann, et al., 2015; Picker, et al., 2009). In other vertebrates, the specification of the NR/RPE domains and the morphogenesis of the optic cup follow a pattern similar to that of the zebrafish, though the relative weight of the different morphogenetic mechanisms may vary among species (Martinez-Morales, et al., 2017).

During the last decades numerous studies have investigated how the RPE and NR domains get genetically specified (Fuhrmann, 2010). Eye identity is established in the anterior neural plate by the early activation of a gene regulatory network (GRN) pivoting on a few key regulators collectively known as eye field transcription factors (EFTF): including Lhx2, Otx2, Pax6, Rx, or Six3 (Zuber, et al., 2003). Upon the influence of inductive signals derived from the lens epithelium (FGFs) or the extraocular mesenchyme (Wnts and BMPs), the eye field network branches into the mutually exclusive developmental programs of the NR and RPE (Cardozo, et al., 2020; Fuhrmann, 2010). Classical genetic experiments performed mainly in mice have identified key nodes of the NR and RPE specification networks. The TF-encoding gene *Vsx2* (previously known as *Chx10*) was identified as the earliest determination gene differentially expressed in NR but not in RPE precursors (Liu, et al., 1994). *Vxs2* plays an essential role in specifying the NR domain by restraining the RPE identity through direct repression of the TF Mitf (Zou and Levine, 2012; Bharti, et al., 2008; Horsford, et al., 2005; Rowan, et al., 2004). Other homeobox regulators, many inherited from the eye field specification network, contribute to the network of NR specifiers. This list includes *Lhx2*, *Sox2*, *Rx*, *Six3* and *Six6* genes, which are required either for NR maintenance or for suppressing RPE identity (Wang, et al., 2016; Roy, et al., 2013). The establishment of the RPE network depends instead on the cooperative activity between Mitf and Otx factors (Lane and Lister, 2012; Martinez-Morales, et al., 2003). Classical experiments in mice showed that inactivation of either *Mitf* or *Otx* genes results in an RPE differentiation failure, with the alternative acquisition of NR molecular and morphological features (Martinez-Morales, et al., 2001; Nguyen and Arnheiter, 2000; Mochii, et al., 1998). More recently, mutation of the main effectors of the Hippo pathway, *Yap* and *Taz*, showed the essential role of these co-regulators in the differentiation of the RPE lineage (Kim, et al., 2016; Miesfeld, et al., 2015).

Despite the identification of these key regulators, the architecture of the NR and RPE specification subnetworks is far from being well understood (Martinez-Morales, 2016). Systematic attempts to reconstruct retinal GRNs using next-generation sequencing (NGS) methods have focused mainly in the differentiation of neuronal types at later stages of development or in epigenetic changes linked to retinal degeneration (Buono and Martinez-Morales, 2020; Wang, et al., 2018; Aldiri, et al., 2017). Very recently, scRNA-seq has been proved a useful tool to characterize cell heterogeneity and infer differentiation trajectories in human tissues and retinal organoids (Collin, et al., 2019; Hu, et al., 2019). However, a comprehensive analysis of the NR and RPE bifurcating networks would require a deeper sequencing coverage approach, while dealing with the limited cell population size at optic cup stages.

In this study, we dealt with previous limitations by applying a combination of RNA-seq and ATAC-seq on sorted NR and RPE populations. Using the zebrafish retina as a model system, we followed transcriptome dynamics and chromatin accessibility changes in both cell populations, as they depart from a common pool of progenitors. The cross-correlation of RNA-seq and ATAC-seq data allowed us to identify activating and repressing cis-regulatory elements (CREs), discover enriched transcription factor (TF) binding motifs, and unveil relevant downstream targets for the main specifiers. Our analyses confirmed previously known TFs as central nodes of the eye GRNs, and more importantly, provide information on their recruitment sequence both in zebrafish and human cells.

Results

1.- Analysis of specification networks in isolated NR and RPE precursors. To examine the bifurcation of the regulatory networks specifying the NR and RPE domains in zebrafish, we focused on the developmental window comprised between stages 16 hpf and 23 hpf (Figure 1 A). Within this period, the optic vesicle transits from a flattened disk in which the lateral and medial layers are still similar in terms of cell shape and volume, to an optic cup stage in which the NR and RPE cells are morphologically differentiated: bottle-shaped for retinal neuroblasts and flat for RPE precursors (Li, et al., 2000). It is important to state that at the end of this window (23 hpf), neurogenesis have not started yet in the retina, neither the RPE display pigmentation (Masai, et al., 2000). We used the transgenic lines *E1_bHLHe40:GFP* and *vsx2.2:GFP* to isolate the NR and RPE populations at 18 and 23 hpf by FACS (see methods; Figure 1B). Both *vsx2* and *bhlhe40* are among the earliest markers reported in zebrafish for the neural retina and RPE domains respectively (Cechmanek and McFarlane, 2017; Barabino, et al., 1997). To isolate early precursors at stage 16 hpf we took advantage of the fact that the *vsx2.2:GFP* transgene is transiently expressed in all progenitors and gets restricted to the NR as progenitors of the medial layer start expressing *bHLHe40* (Cechmanek and McFarlane, 2017; Nicolas-Perez, et al., 2016). Populations isolated by flow cytometry were analysed by RNA-seq, at all stages, and by ATAC-seq at 23 hpf (Figure 1A).

The transcriptomic analysis of the different cell populations highlighted thousands of differentially expressed genes (DEGs) throughout development and between domains (Figure S1A,B; Dataset S1). A multidimensional analysis of these changes revealed that an extensive divergence between the NR and RPE transcriptomes already occurs within the first two hours (16 to 18 hpf) of optic vesicle folding (Figure 1C). Later on, between 18 hpf and 23 hpf, the transcriptome divergence between the NR and RPE progresses, although is more modest within each domain (Figures 1C and S1B).

2.-Chromatin accessibility in NR vs RPE precursors identifies cis-regulatory modules. To gain further insight into the architecture of the optic cup GRNs, we sought to identify differentially open chromatin regions in the morphologically divergent NR and RPE domains. To this end, we performed ATAC-seq experiments from isolated NR and RPE precursors at 23 hpf. A total of 238369 open chromatin regions (OCRs) were detected using this approach. After statistical analysis, a fraction of the peaks was identified as differentially open chromatin regions (DOCRs), corresponding to putative cis-regulatory elements (CRE) that are more active in the NR than in the RPE or vice versa. Approximately 12.6% of all peaks (30172 peaks) were

differentially open with an adjusted p-value < 0.05. This proportion was reduced to 4.8% when the adjusted p-value was lowered to <0.001 (Figure S2A,B; Dataset S2). Regardless the adjusted p-value, we observed a larger number of DOCRs in the RPE than in the NR (18909 vs 11263; adjusted p-value <0.05), and this was also accompanied by a higher average fold change in the RPE than in the NR associated DOCRs (Figure S2B,C). Analysis of the distribution of open chromatin regions in the genome showed no evident differences between NR and RPE for the whole set of OCRs. However, there was a difference in average localization between the entire OCRs and the subsets of DOCRs in the genome. When compared to the distribution of all the OCRs a noticeable depletion of DOCRs located near the promoter regions was evident in both the NR and RPE domains. The proportion of NR peaks near the promoter was 11.91% for OCRs vs 4.6% for DOCRs; whereas for the RPE peaks the proportion was 10.03% for OCRs vs 3.09% for DOCRs (Figure S2D), showing that domain specific cis-regulatory modules tend to occupy more distal positions in the genome. Gene ontology analysis of terms from the category Biological Process enriched in the list of genes associated with DOCRs yield results consistent with the analysed tissue. Genes associated with NR DOCRs were enriched in terms related to nervous system development, neuron differentiation, and eye morphogenesis, whereas genes associated with RPE DOCRs were enriched in terms such as melanocyte differentiation and epithelial differentiation (Figure S3).

Previous reports have shown the importance of correct DNA methylome patterning during human iPSC reprograming towards RPE (Araki, et al., 2019), and have proved a requirement for active demethylation during eye formation in vertebrates (Xu, et al., 2012). Thus, to further characterize the NR and RPE specific DOCRs, we next interrogated their DNA methylome profiles. We utilised base-resolution DNA methylome (WGBS) and hydroxymethylome (TAB-seq) data of zebrafish embryogenesis and adult tissues (Bogdanovic, et al., 2016), to obtain insight into the temporal dynamics of DNA methylation (5mC) over these open chromatin regions. Supervised clustering (k=2) of 5mC patterns separates both datasets into two well-defined clusters (Figure S4A,B). The first cluster consists of distal regulatory regions that initiate active demethylation at ~24hpf, as demonstrated by strong 5-hydroxymethylcytosine (5hmC) enrichment and the progressive loss of 5mC starting at 24hpf, and then are most visible in the adult brain. Importantly, this cluster is highly enriched in H3K27ac, an active enhancer mark, and depleted of H3K4me3, which indicates their identity as bona fide enhancers (Bogdanovic, et al., 2012). The second cluster of both NR and RPE datasets is hypomethylated, depleted of hmC, and enriched in the promoter mark H3K4me3, in line with a CpG island promoters character. To test whether active, tet-protein dependent DNA demethylation is required for the establishment of open chromatin structure at NR and RPE

ATAC-seq peaks, we interrogated chromatin accessibility in wild type and triple tet (tet1/2/3) morphants in whole 24hpf embryos (Bogdanovic, et al., 2016). Data from two independent MO knockdown experiments revealed that their chromatin accessibility signature is significantly (Mann-Whitney-Wilcoxon Test, $P < 2.2e-16$) reduced for NR ATAC-seq peaks in tet morphants when compared to wild type embryos, even when whole embryos were assessed (Figure S4C). This decrease was not visible in RPE peaks, likely because RPE enhancers are active in a much smaller cell population and their chromatin mark is masked by the en masse approach (Figure S4D).

3.- Cross-correlation of RNA-seq and ATAC-seq data reveals activating and repressing CREs. Integration of ATAC-seq and RNA-seq data has been used as a powerful tool to explore the architecture of developmental GRNs (Buono and Martinez-Morales, 2020; Lowe, et al., 2019). By intersecting RNA-seq and ATAC-seq datasets, we observed a substantial overlap between DEGs and genes associated with DOCRs at 23 hpf. Thus, 47% of all DEGs are associated with at least one DOCR (Figure 1D). Specific examples of ATAC-seq and RNA-seq tracks are shown for the *vsx1* and *mitfa* loci, NR and RPE respective markers (Figure 1E,F). The cross-analysis of ATAC-seq and RNA-seq data allowed us to classify DOCRs (putative CREs) in two groups: those correlating with up regulated genes, here termed "activating CREs" and those correlating with silenced genes, here termed "repressing CREs" (Figure 2A,B). Illustrative examples of activating and repressing peaks are provided for the *six3a* and *otx2* loci (Figure S5A,B). When their average distance to the closest Transcription Start Site (TSS) was examined, we observed a significant trend: both for NR and RPE peaks, activating CREs tend to occupy positions closer to the TSS than the repressing regions (Figure S5C). A possible explanation for this difference is the positive bias introduced by proximal promoters, which are typically enriched in binding sites for activators (Maston, et al., 2006). To gain insight into the regulatory logic of the NR and RPE domains, we calculated the number of associated activating and repressing CREs for both the whole DEGs and the subset of differentially expressed TFs. Notably, as expected from their regulatory complexity, the average number of CREs per gene (either activating or repressing) was consistently higher for TFs than for the whole collection of DEGs (Figure 2C). More importantly, in the neural retina TFs associated with activating CREs outnumbered TFs associated with repressing regions (182 vs. 69 respectively), whereas the opposite was observed in the RPE. Indeed, in this tissue we detected a robust repressive cis-regulatory logic, with 173 TFs associated with at least one repressing CREs and only 110 associated with activating ones (Figure 2B,C). This opposite trend, better appreciated in histogram graphs, suggests that the NR specification program is sustained mainly by the activation of

transcriptional regulators, whereas RPE determination seems to require primarily the repression of the NR-specific-TFs (Figure 2D,E).

4.- Distinctive sets of TFs and cytoskeletal components are progressively recruited during NR and RPE specification. To efficiently analyse expression dynamics, as the NR and RPE networks bifurcate, we used a gene clustering approach. We applied both hierarchical and partitioning soft clustering for the classification of genes encoding for TFs or cytoskeletal components. We focused on these two gene categories to examine not only transcriptional lineage specifiers, but also terminal effectors operating in the divergent cell shape changes observed between the NR and RPE domains. Using partitioning soft clustering we established 25 groups of gene expression variation from progenitors towards NR or RPE, for both TFs and cytoskeleton components (Figure S6A,B; Datasets S3 and S4). This approach identified several expression clusters, which are distinctive for each domain and developmental stage: e.g. clusters 1, 13, and 23 for 16hpf progenitors in the TFs category (Figure S6A); or 6, 11, and 21 for 23hpf RPE precursors in the cytoskeletal components' analysis (Figure S6B). Using a hierarchical clustering approach, we could aggregate the small sub-groups into 6 main clusters for TFs and 8 for cytoskeletal components, all of them linked mainly to a specific domain and/or developmental stage (Figures 3 and S7; Datasets S5 and S6). To define precise time windows of expression and to infer regulators order of action, we focused our attention on the identity of TFs belonging to the different large clusters. Significant TFs in cluster #5, corresponding to 16hpf progenitors, include *rx3,* an early eye specifier with a known role in optic vesicle evagination (Rembold, et al., 2006; Loosli, et al., 2001; Mathers, et al., 1997); and *her* factors required to maintain the progenitor neural state (Chapouton, et al., 2011) (Figure 3B). As expected, Cluster #1, corresponding to NR 23 hpf precursors, contains many of the acknowledged retinal specifiers, such as *rx1*, *rx2*, *sox2*, *six3a*, *six3b*, *six6b*, *vsx1*, *vsx2*, *hmx1*, *hmx4* and *lhx2b* (Fuhrmann, 2010), which already increased their expression at 18 hpf (Figure 3B). Surprisingly, known RPE specification genes such as *otx2* and *mitfa* (Martinez-Morales, et al., 2004), included in cluster #4, do not increase their levels significantly in RPE precursors at 18hpf, peaking only later at 23 hpf (Figure 3B). In contrast, TFs included in cluster #3, such as *tead3b*, *tfap2a* and *tfap2c* started to rise, when not peaking, in the RPE at 18 hpf. Notably, some of these TFs, including *tcf12*, *smad6b*, and especially *vgll2b* not only peak at 18 hpf, but also rapidly decrease at 23 hpf (Figure 3B). This observation suggests the existence of two waves of TFs regulating the identity of the RPE domain.

Next, we focused on the content of the hierarchical clusters associated to cytoskeletal components (Figure S7A). While the analysis of the cytoskeletal genes upregulated in PG

returned no significant enrichment for terms annotated in the Cellular Components, similar analysis for cytoskeletal genes in the NR clusters yielded significantly enriched terms associated to microtubules and centrosomes (Figure S7B). This observation is in line with the elongation of the apico-basal axis and the polarization of the microtubules reported for NR precursors in zebrafish (Ivanovitch, et al., 2013). In contrast, the only enriched term in the RPE cytoskeletal clusters was "intermediate filaments" (Figure S7B), consistent with the RPE acquiring a squamous epithelial character.

5.- Motif enrichment analysis suggests redundancy and cooperativity in the NR and RPE networks. To further explore the regulatory logic of the NR and RPE gene networks, we investigated overrepresented motifs within the DOCRs associated with each domain. Motif enrichment analysis of the differentially open regions in the NR identified a highly significant overrepresentation of the homeobox and sox TF binding motifs (Figure 4A; Dataset S7). The core homeobox binding motif (5'- TAATT-3') is shared by TFs from the homeodomain K50 PRD-class; such as *vsx1*, *vsx2*, *rx1*, *rx2*, and *rx3*; LIM-class, such as *lhx2b*; and NKL class; such as *hmx1* and *hmx4* (Figure 4A). All of them are well-known retinal specifiers contained in group #1 of our hierarchical clustering analysis (NR 23 hpf). As they are co-expressed in the retina, it is likely that they cooperate to target a partially overlapping set of cis-regulatory modules. To explore this possibility, we retrieved the individual position weight matrixes (PMW) associated to both homeobox and sox TFs from available databases and used this information to explore their functional synergy in the retinal DOCRs (see Method sections). Two different approaches were used: (i) calculating the co-occurrence rate of binding sites for two different TFs in the same CRE (co-occupancy); and (ii) estimating the percentage of binding sites for TFs located in different CREs but still associated with the same gene (co-regulation) (Figure 4B,C). These analyses predicted an extremely high degree of inter-connectivity within the NR network. Thus, the average co-occurrence rate of two different theoretical binding sites in the same peak (co-occupancy) was 25.6%. This combinatorial activity becomes more pronounced when examining the binding sites in different CREs associated to the same gene (co-regulation), which was in average 46.6%. Indeed, when individual TF binding sites were investigated, the number of co-occupied peaks or co-regulated genes was in general much larger than those instances in which the binding site was found isolated (Figure 4D,E). Even when these are predicted binding sites, our analyses suggest a strong cooperative activity of homeobox and sox factors in driving the NR developmental program.

In the RPE, a motif enrichment analysis of the specific DOCRs also revealed a number of binding motifs that correspond to the TFs identified as up-regulated in the RNA-seq analysis

(Figure 5A; Dataset S8). We found a very significant enrichment for tfap2a and tfap2c binding motifs, in agreement with the early expression of these factors in the RPE. This considerable overrepresentation points to a prominent position of these factors within the regulatory hierarchy, which is in line with their role in the specification of the pigmented tissue (Bassett, et al., 2010). In addition, a significant enrichment for bHLH, tead and otx2 binding motifs was also observed. Importantly, the core bHLH motif (5'-CACGTG-3') is shared by members of the family expressed in RPE, such as *mitfa*, *bHLHE40*, *bHLHE41*, and *tfec* (Figure 5A). As already investigated for the NR TFs, we examined the potential cooperation among these RPE specifiers (Figure 5B-E). Interestingly, whereas the average rate of co-regulated genes in the RPE did not differ much from that found in NR (48% vs 46.6%; Figure 5E), the rate of co-occupied peaks dropped from 25.6% in the NR to 11.7% in the RPE (Figure 5D). This observation suggests that the RPE network is less dependent on TFs cooperativity within the same CRE than the NR network. In such regulatory scenario, each TF may trigger transcriptional sub-programs within a broader developmental network. To explore this possibility, we assessed GO enrichment for the genes associated with RPE DOCRs containing the different TFBS and then we grouped the results using hierarchical clustering (Figure S8). The outcome of this clustering approach supported a branched regulatory scenario. Thus, whereas some TFs (i.e. *tcf12*, *tfap2c*, *otx2*) appeared associated to many GO terms, others were more specific: such is the case of *mitfa* that was associated only to pigment cell differentiation (Figure S8).

Finally, to investigate the regulatory logic of activating and repressing CREs in each domain, we scanned these regions for enriched motifs separately. Strikingly, those enriched motifs ranking higher for the NR activating CREs, such as the homeobox and sox TF binding sites, were also ranking higher for the repressing regions in this tissue (Figure S9; Dataset S9). A very similar scenario was observed for the RPE peaks, which regardless being activating or repressing regions showed tfap2c, tfap2a, bHLH, and otx2 consensus binding sites as top enriched motifs. When GO terms associated with genes linked to these regions were examined, we observed terms linked to eye morphogenesis enriched for both NR activating and RPE repressing regions (Figure S9). This finding reveals context-dependent TFs activity, and points to a common set of genes activated in the NR and repressed in the RPE by antagonistic GRNs. In fact, a detailed analysis of the gene lists associated to NR activating and RPE repressing regions showed that many of the retinal specifiers themselves (including *hmx4*, *lhx9*, *mab21l1*, *nr2f2*, *pax6a*, *pax6b*, *rx2*, *six3a*, or *sox2*) are under the antagonistic regulation of the NR and RPE GRNs (Dataset S10). In addition, the analysis of GO terms enrichment also suggested domain-specific functions for the NR and RPE GRNs. Thus, NR repressing CREs are associated to genes involved

in mesoderm formation, whereas RPE activating regions are link with epidermal differentiation genes.

6.- Desmosomal components are activated during RPE specification. The integration of our ATAC-seq and RNA-seq data allows formulating and testing different hypotheses related to genetic programs controlling the specification of retinal tissues. As an example, we followed up the observation that genes encoding for intermediate filaments are enriched in the RPE domain (Figure S7B). Keratin looping into desmosomal plaques plays a fundamental role in maintaining tissue architecture under mechanical load (Hatzfeld, et al., 2017) (Figure 6A). We thus investigated the transcriptional profile of other desmosomal components as the RPE and NR networks diverge. A detailed analysis revealed that not only keratin genes, but also many desmosomal genes such as *dspa*, *evpla*, *pleca*, or *ppl*, are among the most upregulated genes (i.e. highest fold change) in committed RPE cells (Figure 6B). Many of these genes increased their expression already at 18 hpf, when RPE cells start to differentiate morphologically, but hours before the tissue acquires pigmentation (Figure 6B). To gain insight into the regulation of these cytoskeletal components, we performed a motif enrichment analysis of the subset of DOCRs associated with keratins and other desmosomal genes upregulated in the RPE (Dataset S11). This analysis revealed Tead motifs as top ranked overrepresented binding sites in the CREs regulating desmosomal genes (Figure 6C). When compared to that of the entire set of DOCRs, the average number of motifs per peak was 4.5 fold higher for those CREs specifically associated with components of the desmosome machinery (1.35 vs 0.3). This observation suggests a key role for Tead family proteins in the regulation of the intermediate filament cytoskeleton. To test this hypothesis, we took advantage of available zebrafish double mutants for the main Tead coactivators *yap* and *taz*, which have been reported to display RPE differentiation defects (Miesfeld, et al., 2015). In agreement with our observations, the expression of keratins (*krt4* and *krt8*) as well as pigmentation genes (*tyr* and *tyrp1b*) were significantly reduced in *yap -/- taz -/-* double mutant head tissue at 18 hpf as determine by RT-qPCR (Figure 6D). As a control, the expression of the NR markers *vsx2* and *six3a* was unaffected in the mutant tissue. These results further support a role for Tead factors in the transcriptional regulation of desmosomal genes at the RPE.

The high efficiency of the CRISPR/Cas9 technology allows performing F0 mutagenesis screens in zebrafish (Shankaran, et al., 2017). We took advantage of this technology to test functionally some of the genes identified in our analyses. A total of 21 genes, selected on the bases of their expression profile, CRE composition and/or dynamics, associated GO terms, and number of paralogues, were investigated in an F0 pilot screen (Figure S10). A substantial proportion of

the tested genes (i.e. 12 out of 21 sgRNA combinations tested) displayed eye malformations, such as microphthalmia, eye fissure closure defects and/or hypopigmentation, thus confirming that many of the candidate genes play an essential role as components of the eye GRNs (Figure S10; Dataset S12). Interestingly, among the sgRNA combinations tested, the injection of those directed against the structural desmosomal components *dspa/b*, the Tead regulators *vgll2a/b*, and particularly the cocktail *tead1a/b* + *tead3a/b* resulted in a significant proportion (30.4%, 40.6%, and 80% respectively) of the F0 embryos displaying reduced eye size and severe hypopigmentation (Figures 6E and S10). These results further indicate that both desmosomal assembly and Tead activity are required during RPE differentiation.

7.- Gene expression analysis during hiPSCs-to-RPE differentiation. A second important finding derived from our gene expression clustering analysis was the identification of two waves of transcriptional regulators during the specification of the RPE in zebrafish. Understanding this TF recruitment sequence in humans may have important basic and translational applications. We thus asked whether the same *consecutio temporum* identified in zebrafish was conserved in human iPSCs differentiating to RPE. During the first four weeks of differentiation in culture from hiPSCs (see methods), pluripotent stem cells progressively change their morphology to a cobblestone appearance, acquiring a light pigmentation at the end of the fourth week (Figure 7A). Using RT-qPCR, we tested mRNA levels for a total of 26 RPE genes including genes encoding for stemness markers (*NANOG* and *OCT4*); mature RPE markers (i.e. *CRALBP, RPE65* and *TYR*); known RPE specifiers activated in the second wave of gene expression (i.e. *BHLHE40, MITF*, *OTX2*, and *TFEC*); desmosomal components (i.e. *DSP*, *EVPL*, *KRT4*, *KRT5*, *KRT8*); and TFs and signalling molecules activated in the first wave of gene expression (NOTCH1, NOTCH2, NOTCH3, SMAD6, TCF12, TEAD1, TEAD2, TEAD3, TEAD4, TFAP2A, TFAP2C, and VGLL2) (Figure 7B; Dataset S13). This analysis showed that most of the human genes orthologous to the zebrafish early specifiers cluster together according to their expression profiles, and reach maximal expression within the first weeks of culture. This was the case for *SMAD6*, *TEAD1*, *TEAD3*, *TFAP2A*, *TFAP2C*, and particularly for *VGLL2 and TEAD2,* the expression of which peaked very transiently. Similarly, the human orthologs of most of the genes that in zebrafish were identified as activated in a second transcriptional wave (including *MITF*, *BHLHE40*, *TFEC, CRALBP, RPE65* and *TYR*) also clustered together reaching a maximum expression level in the fourth week of culture. These findings indicated that despite the very different time scales of the developmental programs in these far-related vertebrate species, hours in zebrafish and days in humans, their RPE regulatory networks share a common logic of two-waves in the TFs recruitment sequence.

Discussion

A system level approach to explore GRNs architecture. Cell differentiation, as classically illustrated by Waddington's epigenetic landscapes (Waddington, 1966), often entails sequential binary decisions that progressively restrict cell competence. At a genetic level this translates in the branching of "mother" gene regulatory networks into child sub-networks. Here we have used a double RNA-seq/ATAC-seq approach to zoom into the bifurcation of the NR/RPE developmental programs, which will give rise to specialized neurons and pigmented squamous cells respectively. The transient nature of this differentiation process, as well as the limiting size of the cell populations involved (Kwan, et al., 2012), have hindered any systemic approach to investigate the architecture of the specification networks as they branch. The high sensitivity and deep sequencing coverage of the RNA-seq and ATAC-seq methods (Buenrostro, et al., 2013), combined with the isolation of NR and RPE precursors by FACS allowed us to overcome previous constrains. Our approach permitted not only the detection of transcriptomic variations and active cis-regulatory modules, but also helped to define hierarchical relationships among the core components of the network. In our analyses we have used a proximity method to associate genes and open chromatin regions. Previous studies have confirmed that this link to the adjacent gene is correct in 90% of the cases (Yoshida, et al., 2019). Therefore, although this estimate should be taken cautiously when individual cis-regulatory elements are examined, it is a valid approach when conclusions are driven from a genome-wide analysis. In this study we identified approximately 30000 chromatin regions differentially open (DOCRs) between the NR and RPE domains. The examination of their methylation status, as well as their overlap with H3K27ac and H3K4me3 marks indicates that most of these regions (≈95%) correspond to initially inactive enhancers that get progressively demethylated and activated at the phylotypic period, a behaviour reported for developmental genes in general (Bogdanovic, et al., 2016). A much smaller cluster (≈ 5%) matches to hypomethylated constitutively active promoter regions (Figures S2 and S4). It is therefore likely that most of the regions here identified as DOCRs correspond to cis-regulatory elements (CREs) active during the segregation of the NR/RPE networks.

The NR as a default program. Our data, together with previous observations by others, are consistent with the hypothesis that the neural retina is the default state of the optic vesicle precursors. Morphologically, optic vesicle precursors either at the medial or the lateral layers share a similar neuroepithelial character: i.e. elongated cells polarized along the apico-basal axis arranged in a pseudostratified epithelium (Ivanovitch, et al., 2013; Kwan, et al., 2012).

During optic cup morphogenesis, NR precursors retain this neuroepithelial morphology, whereas differentiating RPE cells undergo profound cell shape changes as they progressively flatten into a squamous epithelium (Moreno-Marmol, et al., 2018). These observations correlate in our analyses with a larger number of differentially upregulated genes and differentially opened chromatin regions at the RPE (Figure S2). We have shown that repressing CREs associated with TFs dominate in the RPE network, whereas activating elements are more abundant in the NR program (Figure 2). Furthermore, the motifs analysis of activating elements in the NR and repressing elements in the RPE indicates that they are antagonistically regulating a similar set of genes (Figure S9). It is therefore very likely that the acquisition of the RPE fate requires the global repression of the NR program. However, although the role of *Vsx2* in the suppression of RPE identity via direct *Mitf* repression has been well documented (Zou and Levine, 2012; Horsford, et al., 2005; Rowan, et al., 2004), the mechanisms underlining the repressive activity of the RPE network are much less understood (Fuhrmann, et al., 2000). Interestingly, *Pax6* seems to cooperate both with NR and RPE specifiers acting as a balancing factor between the two networks depending on the cellular context (Raviv, et al., 2014; Bharti, et al., 2012). The complexity of the transcriptional regulatory logic is often exemplified by its context dependence, which may happen at very different levels. As illustrated by a recent report in *Drosophila*, cis-regulatory modules can play a dual role as enhancers or silencers in different cellular environments (Gisselbrecht, et al., 2020). Our motif analysis also points to the same set of TFs acting as repressors or activators within the same tissue, suggesting their activity is conditioned by their chromatin environment. This is in line with previous work showing that TFs have diverse regulatory function depending on the enhancer context (Stampfel, et al., 2015).

GRNs robustness. One of our main observations is that NR and RPE networks are remarkably robust. This is the case particularly in teleosts, in which whole genome duplication resulted in additional copies of eye specifiers. In zebrafish, several homeobox TFs expressed in the neural retina converge on the 5'- TAATT-3' motif, including vsx1, vsx2, rx1, rx2, rx3, lhx2b, lhx9, hmx1 and hmx4. Our motif discovery predictions suggest these proteins, together with sox factors, may coregulate the same genes and cooperate within the same CREs, although these are theoretical predictions that may require validation by direct ChIP-seq studies. This approach though is challenging in our experimental setting due to the limited starting material and the reduced availability of ChIP grade antibodies in zebrafish. Yet, we believe that considered in bulk these are informative predictions, as they are restricted to differentially open chromatin regions and differentially expressed TFs. Supporting a co-regulation, the mutation of many

retinal homeobox TFs, such as *rx2* in medaka (Reinhardt, et al., 2015) or *lhx2* in zebrafish (Seth, et al., 2006) does not compromise severely the identity of the tissue. In the RPE, redundancy pertains mainly the core bHLH motif (5'-CACGTG-3'), which is shared by mitfa, bHLHE40, bHLHE41, and tfec. In agreement with this, the simultaneous elimination of *mitfa* and *mitfb* has no apparent consequences on RPE specification or pigmentation in zebrafish (Lane and Lister, 2012). All these observations further point to a broad redundancy and cooperativity among the main nodes of the eye specification networks.

Coupling of transcriptional and cell shape changes in the RPE. As expected, the cell shape changes associated to the acquisition of the NR and RPE identities have a clear reflection at the transcriptomic level. The differentiation of the NR entails the activation of cytoskeletal components associated to microtubules polarization, in agreement with the neuroepithelial character of this tissue (Ivanovitch, et al., 2013). In contrast, here we show that a significant number of keratins and other desmosomal genes get recruited in the presumptive RPE hours before this tissue acquires its distinctive pigmentation. Although keratins and plaque proteins have been reported as RPE markers in several vertebrate species (Owaribe, et al., 1988), their contribution to the morphogenetic program of this tissue remains unexplored. Given the role of desmosomal plaques in conferring resistance to mechanical load (Hatzfeld, et al., 2017), it is tempting to hypothesize that the activation of these cytoskeletal genes may be linked to RPE precursors adapting their cytoskeleton to increased tissue tension as they flatten. Motif enrichment analysis of CREs linked to desmosomal genes, together with the genetic evidence here provided by *yap -/-*; *taz -/-* double mutants and Tead crispants, strongly support a role for Tead TFs in the activation of keratin genes at the RPE. Indeed, several keratin (e.g. *krt5*, *krt8*, and *krt97*), and desmosomal genes (e.g. *evpla*, *pleca*, *plecb*, and *ppl*) have been identified as targets for Yap/Tead complexes by ChIP-seq studies in mammalian cells (Estaras, et al., 2017; Zanconato, et al., 2015; Lian, et al., 2010), as well as DamID-seq studies in zebrafish embryos (Vazquez-Marin, et al., 2019). Thus, the direct transcriptional regulation of keratins and plaque genes by Yap/Tead complexes seems a conserved theme across different tissues and vertebrate species.

Two-waves of TFs recruitment during RPE specification. Arguably one of the most notable observations in this study is the relatively late peak of expression of genes such as *mitfa* or *otx2*, which were previously considered among the earliest RPE specifiers (Martinez-Morales, et al., 2004). Their delayed peak of expression occurs after the specification and flattening of RPE precursors have commenced, and rather coincides with the onset of pigmentation in the

tissue at 23hpf. This agrees with previous studies indicating a cooperative role for Otx and Mitf in the direct activation of the melanogenic gene battery (Martinez-Morales, et al., 2003). A detailed expression study in zebrafish also reported two separated phases for RPE specification, the first of which is *mitf* independent (Cechmanek and McFarlane, 2017). Here we show that the early phase of specification entails the recruitment of a different set of transcriptional regulators, including *smad6b*, *tead1b*, *tead3b*, *tfap2a*, *tfap2c*, *tcf12*, and *vgll2b*. Some of these TFs, or their cofactors, act as upstream regulators of *Otx* and *Mitf*, such is the case for *Tcf* (Westenskow, et al., 2009) and Tead-cofactors *Yap* and *Taz* (Kim, et al., 2016; Miesfeld, et al., 2015). Others, such as *tfap2a*, play an important role in RPE specification, although their hierarchical role within the RPE GRN remains unclear (Bassett, et al., 2010). Finally, here we also describe a similar sequence of TF recruitment in human differentiating RPE cells. Thus, despite the different morphology of the cells (i.e. cuboidal in humans and squamous in zebrafish) and the large evolutionary distance (≈450 mya), the regulatory logic that specifies the RPE seems largely conserved across vertebrates.

Conclusions. Our data shed light on the bifurcation of the NR and RPE programs not only in zebrafish but in other vertebrate species. The obtained results may be relevant to identify novel causative genes for eye hereditary diseases, as mutations in many nodes of the eye GRNs result in congenital eye malformations (Fitzpatrick and van Heyningen, 2005). Importantly, our findings uncover an unanticipated regulatory logic within the RPE specification network. This provides critical information to improve hiPSCs-to-RPE differentiation protocols, a key step in cell replacement strategies for retinal degenerative diseases.

Methods

Fish maintenance

The zebrafish (Danio rerio) AB/Tübingen (AB/TU) wild-type strains, the transgenic lines tg(vsx2.2:GFP-caax) (Nicolas-Perez, et al., 2016) and tg(E1_bHLHe40:GFP) (Moreno-Marmol et al., 2020) and the mutant strain yap +/- taz +/- (Miesfeld, et al., 2015) were maintained and breed under previously described experimental conditions (Westerfield, 2000). All animal experiments were carried out according to the guidelines of our Institutional Animal Ethics Committee.

Cell cytometry

Zebrafish cells were dissociated and prepared for FACS as previously described (Manoli and Driever, 2012). PG at 16 hpf, NR at 18 and 23 hpf were isolated from dissected heads of the tg(*vsx2.2:caax-GFP*). RPE at 18 and 23 hpf were isolated from whole tg(E1_*bHLHE40*:GFP) embryos. A FACSAriaTM Fusion flow cytometer was used to recollect only the GFP+ cells. GFP+ cells were isolated directly in Trizol for RNA extraction, or in ATAC-seq tagmentation buffer for open chromatin detection.

RNA extraction

Total RNA was extracted using 750 ul TRIzol LS (Invitrogen) following manufacturer's protocol. Possible DNA contamination was eliminated treating the RNA samples with with TURBO DNAse-free (Ambion). Concentration of the RNA samples was evaluated by Qubit (Thermo Fisher), and then the samples were used for subsequent applications.

qPCR

cDNA retrotranscription and qPCR were performed as described (Vazquez-Marin, et al., 2019). For primer sequences, see Tables S1 and S2. *HPRT1* and *GAPDH* were used as housekeeping genes for human samples, whereas *eef1a1l1* was used for zebrafish samples.

RNA-seq analyses.

RNA was extracted from sorted cells and then treated with DNAse as described above. rRNAs were eliminated from the samples with Ribo-Zero® rRNA Removal Kit (Illumina) prior library preparation. Samples were sequenced in SEx125bps or PEx125bps reads with an Illumina Hiseq 2500. We obtained at least 35 M reads from the sequencing of each library. Three biological replicates were used for each analyzed condition. Reads were aligned to the danRer10

zebrafish genome assembly using Tophat v2.1.0. Transcript abundance was estimated with Cufflinks v2.2.1. Differential gene expression analysis was performed using Cuffdiff v2.2.1, setting a corrected p value < 0.05 as the cutoff for statistical significance of the differential expression. Multidimensional Scaling Analysis (MDS) was performed using the function MDSplot of the CummeRbund package in R 3.6.1. Soft clustering of time-series gene expression data was done for all the transcripts with a variance among the five conditions ≥ 3 using the R package Mfuzz with a m = 1.5 (Kumar and M, 2007). The TF transcript subset was extracted from the total list of genes using the tool "Classification System" of PANTHER (Thomas, et al., 2003) filtering for the protein class "transcription factors" (PC00218). Some TFs not present in the database for an annotation issues (i.e. mitfa, vsx1, vsx2, rx1, rx2, rx3, lhx2b, hmx1, hmx4, sox21a) were added manually. The cytoskeleton component subset was obtained retrieving all the genes belonging to GO term "cytoskeleton" (GO:0005856), including all the child and further descendant GO terms, with biomaRt. All the heatmaps representing transcriptomic variations were plotted with the R package *pheatmap* using exclusively the transcripts that resulted to be differentially expressed from the comparison between at least two of our experimental conditions TFs and cytoskeleton components were filtered using the same methodology used for Mfuzz clustering. Gene ontology analysis was performed with the online tool GOrilla (Eden, et al., 2009) or Panther (Thomas, et al., 2003) using two unranked lists of genes (target and background lists).

ATAC-seq

ATAC-seq was performed starting from 5000 sorted cells using a FAST-ATAC protocol previously described (Corces, et al., 2016). All the libraries were sequenced 2x50 bp with an Illumina Hiseq 2500 platform. We obtained at least 100 M reads from the sequencing of each library. For data comparison, we used two biological replicates for each condition. Reads were aligned to the danRer10 zebrafish genome and differential chromatin accessibility was calculated as reported (Magri, et al., 2019). All chromatin regions reporting a differential accessibility with an adjusted p-value < 0.05 were considered as differentially open chromatin regions (DOCRs). All the DOCRS have been associated with genes using the online tool GREAT (McLean, et al., 2010) with the option "basal plus extension". Gene ontology analysis of all the genes associated with DOCRs was also performed with GREAT. De novo motif enrichment of TF binding sequences in the sets of DOCRs was performed using HOMER (Heinz, et al., 2010). Top enriched TF PWMs from the HOMER results and PWMs from JASPAR database (Fornes, et al., 2020) were used as input for the online tool FIMO (Grant, et al., 2011) to assess the exact TFBS genome position in the DOCRs. Before estimating the rate of TF co-occupancy in same peak

among the binding motifs for the different TFs, all the binding motif sequences overlapping for more than 3 bps were eliminated, keeping only the TF binding sequences with the lowest p-value.

Activating/repressing cis-regulatory element configuration

For this analysis only the DEGs and the DORCs that could be associated to each other were taken into account. The log2FoldChange values of transcript expression and chromatin accessibility of NR and RPE at 23 hpf were used to discriminate four clusters of activating and repressing CREs. Gene Ontology enrichment analysis of the genes belonging to the different clusters was performed with FishEnrichr (Kuleshov, et al., 2016).

CRE 5mC analysis

Whole genome bisulfite sequencing and TET-assisted bisulfite sequencing (Bogdanovic, et al., 2016) data was trimmed with Trimmomatic software (Bolger, et al., 2014) and mapped onto the danRer10 genome assembly using WALT (Chen, et al., 2016) keeping only the reads mapping to a unique genomic location. Duplicates were removed with sambamba (Tarasov, et al., 2015) and DNA methylation levels were calculated using MethylDackel (https://github.com/dpryan79/MethylDackel). All heatmaps were made using Deeptools (Ramirez, et al., 2014) and Gene ontology enrichments were calculated with GREAT (McLean, et al., 2010). ATAC-seq reads were counted using BEDtools (Quinlan and Hall, 2010) and statistics were performed in R.

CRISPR/Cas9 *F0 screening*

All the sgRNAs were designed using the online tool CRISPscan (https://www.crisprscan.org/) (Moreno-Mateos, et al., 2015) and synthetized following described protocols (Vejnar, et al., 2016). All the sgRNAs were selected to target the first half of the CDS in exons resulting actually expressed in the eye tissues from our RNA-seq data (trying to avoid the first exon to prevent the usage of an alternative start codon that would produce a possibly functional protein), with an efficiency score > 58 and no predicted off-targets (Table S3). Two different sgRNAs were used together to target the same gene. The sgRNAs were injected in the zebrafish yolk at 1-cell stage at final concentration of 80 ng/µl together with the Cas9 endonuclease at a concentration of 300 ng/µl. 1 nl of the mixture was injected in each embryo. In the case a target gene had a close paralogue, the sgRNAs targeting both of the paralogues were injected at the same time, adjusting the final concentration of the sgRNA-Cas9 mixture. Lethality, phenotypic features and penetrance were assessed at 24 and 48 hpf.

Cell culture

Human hiPSCs were obtained, after informed consent, from peripheral blood monocytes by cell reprogramming using a non-integrative Sendai virus vector as described (Garcia Delgado, et al., 2019). Cells were maintained in feeder-free adherent conditions onto Matrigel-covered plates in standard incubation at 37ºC, 5% CO_2, 20% O_2. hiPSCs were fed every two days with mTser1 serum-free culture medium and passaged every 5-7 days depending on confluency. Dispase was used for gentle dissociation for passage. A 6-well plate with undifferentiated, well-grown hiPSCs was the starting point of the experiment (Day 0). To induce RPE differentiation culture medium was changed to the following: KO DMEM, KSR 15%, Glutamax 2 mM, non-essential aminoacids 0.1 mM, β-mercaptoethanol 0.23 mM, Peniciline/streptomycin. Differentiating cells were harvested directly in Trizol LS (Invitrogen) at day 0, and weeks 1, 2, 3 and 4 for gene expression studies.

Acknowledgements

We thank Lázaro Centanin, Juan Tena, Miguel Moreno-Mateos, Joaquin Letelier, Marta Magri, and Santiago Negueruela for their excellent scientific advice and critical input. Corin Díaz and Katherina García for their technical assistance with FACS. This work is supported by the following grants: (I) To JRMM: From the Spanish Ministry of Science, Innovation and Universities (MICINN): BFU2017-86339P with FEDER funds, and MDM-2016-0687. (II) To OB Australian Research Council (ARC) Discovery Project (DP190103852). (III) To FJDC: Andalusian Ministry of Health, Equality and Social Policies (PI-0099-2018). (IV) To PB: BFU2016-75412-R with FEDER funds; PCIN-2015-176-C02-01/ERA-Net Neuron ImproVision, and a CBMSO Institutional grant from the Fundación Ramón Areces. (V) To both JRMM and PB: BFU2016-81887-REDT, as well as Fundación Ramón Areces-2016 (Supporting LB).

Author contribution.

LB conducted most experiments, performed bioinformatic analyses, and had a main contribution in figures and manuscript edition. SN, TMM, and RP contributed to protocols refining, FACS experiments, and fish maintenance. BdC and FJDC carried out hiPSCs experiments together with LB. OB performed 5mC analysis and contributed to manuscript edition. JRMM conceived the project and assisted LB in data analysis. The manuscript was edited and written by JRMM and PB, who co-supervised all the work.

Conflict of Interest.

The authors declare that they have no conflict of interest

Availability of Data

Datasets supporting the conclusions of this article are available in the Gene Expression Omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/geo) under the following accession numbers: RNA-seq (GSE150346) and ATAC-seq (GSE150189).

Figure Legends: (max 300 words)

Figure 1: ***Experimental setup and raw data*** (A) Schematic representation of optic vesicle morphogenesis from undifferentiated retinal progenitors cells (PG, blue), at 16 hpf, to segregation and differentiation of the neural retina (NR, purple) and pigmented epithelium (RPE, orange) domains at 18 and 23 hpf. NR and RPE populations isolated by flow cytometry were analysed by RNA-seq and ATAC-seq at the stages indicated. (B) Stage 23 hpf confocal sections from the zebrafish transgenic lines used to mark and isolate the RPE (*E1_bHLHe40:GFP*) and NR (*vsx2.2:GFP*) populations. (C) Multidimensional scaling analysis of the RNA-seq data showing a progressive transcriptomic divergence between the NR and RPE domains. (D) Percentage of overlapping between differentially expressed genes (DEGs) and genes associated with differentially opened chromatin regions (DOCRs). (E-F) Overview of ATAC-seq and RNA-seq tracks (UCSC browser) for representative NR (*vsx1*; E) and RPE (*mitfa*; F) specific genes. Solid bars on the top indicate DOCRs. If purple, the DOCR is more accessible in the NR. On the contrary, if orange, the DOCR is more accessible in the RPE. In the depicted case all the DOCRs are accompanied by an increased transcription of the associated gene in the corresponding tissue.

Figure 2: ***CRE configuration in the NR and RPE domains***. (A) Schematic representation of the functional relationship between DOCRs, either activating (green) or repressing (red), and their associated DEGs. (B) The graph illustrates the correlation between the levels of differentially expressed TFs (log2FC) and the accessibility of their associated DOCRs (log2). (C) The table summarizes the number of activating or repressing CREs associated with either to all the DEGs or only with the differentially expressed TFs. (D-E) Histograms correlating the number of TFs associated with activating or repressing CREs to the number of peaks per gene, for both the NR (D) and RPE (E) domains.

Figure 3: ***Hierarchical clustering of TFs gene expression variations during optic cup development.*** (A) Hierarchical clustering output shows TFs expression trends in the distinct domains and stages. Gene expression values, normalized by row, are indicated with a red to blue graded colour. Note that most clusters comprise a particular domain and developmental stage. (B) Relative transcript level changes for significant TFs (indicated by dotted lines) within each cluster.

Figure 4: ***NR motif enrichment analysis***. (A) Representative TF binding motifs enriched in NR DOCRs as identified by HOMER. The binding motif similarity among neural retina TFs of the homeobox family is indicated (http://jaspar.genereg.net). (B) Circoplot illustrating the co-occupancy rate of TFBS in the same DOCRs for the main TFs identified in the motif enrichment

analysis (C) Circoplot illustrating the degree of co-regulation between TFs regulating the same gene through different DOCRs. (D) Number of CREs containing the main TFs identified in the motif enrichment analysis, classified according to their co-occupancy (E) Number of genes associated to CREs containing the main TFs identified in the motif enrichment analysis, classified according to their co-regulation.

Figure 5: ***RPE motif enrichment analysis***. (A) Representative TF binding motifs enriched in RPE DOCRs as identified by HOMER. Analysis of the binding motif similarity among TFs of the bHLH family (http://jaspar.genereg.net). (B) Circoplot illustrating the co-occupancy rate of TFBS in the same DOCRs for the main TFs identified in the motif enrichment analysis (C) Circoplot illustrating the degree of co-regulation between TFs regulating the same gene through different DOCRs. (D) Average percentage of co-occupancy in the same DOCR for two different TF in the RPE and NR. (E) Average percentage of two different TF regulating the same gene through different DOCRs in the RPE and NR.

Figure 6: ***Regulation of desmosomal components during RPE specification.*** (A) Schematic representation of a desmosome junction. (B) Intermediate filament and desmosomal genes expression variations during optic cup development. Expression values are reported as log(FPKM) (C) Motif enrichment analysis of the DOCRs associated with genes encoding intermediate filament or desmosome components. (D) mRNA levels of keratin genes as well as RPE and NR markers as determined by RT-qPCR in wild type and *yap -/- taz -/-* double mutant zebrafish samples (dissected heads) at 18 hpf. Significant differences are indicated (n=3; T-test; ***= $p<0.001$). (E) Representative stereo microscope images of zebrafish embryo heads at 48 hpf. Magnification Bar = 0.5 mm): wild type and embryos injected with Cas9 (300 ng/µl) together with the following sgRNAs (80 ng/µl) combinations: *vgll2a* and *vgll2b* (vgll2)*; tead1a* and *tead1b* (tead1) and *dspa and dspb* (dsp). Note the reduced eye size and RPE hypopigmentation in the crispants.

Figure 7: Gene expression during hiPSCs-to-RPE differentiation. (A) Bright field microscopy images (10-x and 20x) of hiPSCs before (Day 0) and during their differentiation towards RPE. Note the progressive acquisition of epithelial morphology and pigmentation. (B) Hierarchical clustered heatmap showing gene expression level variations, as determined by RT-qPCR during the differentiation towards RPE from hiPSCs. Note the conservation of the two clusters (early and late transcriptional waves) identified in zebrafish.

Supplementary Information. Figure Legends:

Figure S1: Eye domain transcriptome variations during optic cup morphogenesis. (A) Volcano plots illustrating the transcriptome variations during eye morphogenesis. Each dot corresponds to a gene. Black dots indicate not significant variations, whereas coloured dots point out significant expression variations among domains and developmental stages. (B) Table summarizing the number of DEGs (upregulated or downregulated) in each one of the conditions.

Figure S2: Analysis of differentially opened chromatin regions (DOCRs) during optic cup morphogenesis. (A) Volcano plot illustrating the chromatin accessibility changes during eye morphogenesis. Each dot corresponds to a peak (i.e. open chromatin region). Black dots indicate not significant variations; colour shades chromatin accessibility changes with different ranges of adjusted p-value (darker: $p<0.05$; medium= $p<0.01$, lighter= $p<0.001$). (B) Table including the number of peaks significantly more or less accessible between the two conditions. (C) Frequency histogram showing the distribution of the DOCRs in relation to their accessibility fold change. (D) Pie charts displaying the percentage of peaks falling in distinct regions of zebrafish genome. Top: genome distribution of the whole set of open chromatin regions (OCRs) identified by ATAC-seq in each condition. Bottom: genome distribution of the DOCRs.

Figure S3: Gene ontology enrichment of the genes associated with DOCRs. Bar chart showing GO terms for biological processes enriched in the genes associated to DOCRs. Bar length is proportional to enrichment significance.

Figure S4: DNA methylation profiles of NR and RPE peaks. Clustering of NR (A) and RPE (B) ATAC-seq peaks according to 5mC pattern across developmental stages and in adult brain (left columns), as well as 5hmC, K4me3, and K27ac profiles at 24 hpf (right columns). Note that the signature of cluster #1 corresponds to active enhancers, whereas that of cluster #2 corresponds to hypomethylated promoter regions. (C) Chromatin accessibility quantification for NR and RPE peaks in wild type and triple tet (tet1/2/3) morphants. Analysis was performed in whole 24hpf embryos. Statistical significance was evaluated using a Mann-Whitney-Wilcoxon Test, $P < 2.2e-16$. (D) Chromatin accessibility levels for NR and RPE ATAC-seq peaks in wild type and and triple tet (tet1/2/3) zebrafish morphants.

Figure S5: Activating and repressing CREs localization. (A) Schematic representation of the functional relations among DOCRs, either activating (green) or repressing (red), and their associated DEGs. (B) Illustrative examples of ATAC-seq and RNA-seq tracks (UCSC browser) for representative NR (*six3*; E) and RPE (*otx2*; F) genes. Bars on the top indicate DOCRs. If purple,

the DOCR is more accessible in NR; if orange, the DOCR is more accessible in RPE. Some of the more accessible DOCRs are accompanied by a decreased transcription of the associated gene in the corresponding tissue. (C) Average distance to the TSS of all activating and repressing DOCRs for the NR and RPE domains (see Figure 2C). Note that activating CREs are significantly (T-test) closer to the TSS than repressing regions in both domains.

Figure S6: Partitioning clustering of gene expression variations during optic cup development. Partitioning clustering output (k= 25) showing the expression trends in the distinct domains and stages examined for two classes of genes: TFs (A) and cytoskeleton components (B). Stages and domains examined are linearly represented with a colour code and numbered from 1 (NR 23hpf) to 5 (RPE 23hpf), following the lineage trajectory.

Figure S7: Hierarchical clustering of gene expression variations during optic cup development for cytoskeletal components. (A) Hierarchical clustering output showing the expression trends in distinct domains and stages. Gene expression values, normalized by row, are indicated with a red to blue graded colour. Note that all stages and domains are represented by at least one gene cluster (B) Cellular component GO enrichment for cytoskeletal genes belonging to NR or RPE clusters.

Figure S8: GO terms associated with RPE TFs. Hierarchical clustering of the main RPE TFs according to the enriched GO terms (Biological process) associated to their putative downstream genes (genes associated with RPE DOCRs that contain each TFBS).

Figure S9: ***Motif enrichment analysis of activating and repressing CREs***. Representative TF motifs enriched in activating (left column) and repressing (right column) CREs in both the NR (A) and RPE (B) domains. Analysis of GO terms enrichment for genes associated with each set of elements is indicated.

Figure S10: CRISPR CRISPR/Cas9 F0 screening. (A) Eye phenotypes resulting from the injection of Cas9 protein and sgRNAs for the candidate genes. (B) Percentage of embryos showing an impaired eye phenotype upon the injection of Cas9/sgRNAs complex (n= total number of injected embryos). The % of lethality upon injection is also shown. See also Supplementary data set 12.

References

Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., Shirinifard, A., Thiagarajan, S., Sablauer, A., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. Neuron 94, 550-568 e10.

Araki, H., Miura, F., Watanabe, A., Morinaga, C., Kitaoka, F., Kitano, Y., Sakai, N., Shibata, Y., Terada, M., Goto, S., et al. (2019). Base-Resolution Methylome of Retinal Pigment Epithelial Cells Used in the First Trial of Human Induced Pluripotent Stem Cell-Based Autologous Transplantation. Stem Cell Reports 13, 761-774.

Arendt, D., and Wittbrodt, J. (2001). Reconstructing the eyes of Urbilateria. Philos Trans R Soc Lond B Biol Sci 356, 1545-63.

Barabino, S.M., Spada, F., Cotelli, F., and Boncinelli, E. (1997). Inactivation of the zebrafish homologue of Chx10 by antisense oligonucleotides causes eye malformations similar to the ocular retardation phenotype. Mech Dev 63, 133-43.

Bassett, E.A., Williams, T., Zacharias, A.L., Gage, P.J., Fuhrmann, S., and West-Mays, J.A. (2010). AP-2alpha knockout mice exhibit optic cup patterning defects and failure of optic stalk morphogenesis. Hum Mol Genet 19, 1791-804.

Bharti, K., Gasper, M., Ou, J., Brucato, M., Clore-Gronenborn, K., Pickel, J., and Arnheiter, H. (2012). A regulatory loop involving PAX6, MITF, and WNT signaling controls retinal pigment epithelium development. PLoS Genet 8, e1002757.

Bharti, K., Liu, W., Csermely, T., Bertuzzi, S., and Arnheiter, H. (2008). Alternative promoter use in eye development: the complex role and regulation of the transcription factor MITF. Development 135, 1169-78.

Bogdanovic, O., Fernandez-Minan, A., Tena, J.J., de la Calle-Mustienes, E., Hidalgo, C., van Kruysbergen, I., van Heeringen, S.J., Veenstra, G.J., and Gomez-Skarmeta, J.L. (2012). Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome Res 22, 2043-53.

Bogdanovic, O., Smits, A.H., de la Calle Mustienes, E., Tena, J.J., Ford, E., Williams, R., Senanayake, U., Schultz, M.D., Hontelez, S., van Kruijsbergen, I., et al. (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. Nat Genet 48, 417-26.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-20.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10, 1213-8.

Buono, L., and Martinez-Morales, J.R. (2020). Retina Development in Vertebrates: Systems Biology Approaches to Understanding Genetic Programs: On the Contribution of Next-Generation Sequencing Methods to the Characterization of the Regulatory Networks Controlling Vertebrate Eye Development. Bioessays 42, e1900187.

Cardozo, M.J., Almuedo-Castillo, M., and Bovolenta, P. (2020). Patterning the Vertebrate Retina with Morphogenetic Signaling Pathways. Neuroscientist 26, 185-196.

Cechmanek, P.B., and McFarlane, S. (2017). Retinal pigment epithelium expansion around the neural retina occurs in two separate phases with distinct mechanisms. Dev Dyn 246, 598-609.

Chapouton, P., Webb, K.J., Stigloher, C., Alunni, A., Adolf, B., Hesl, B., Topp, S., Kremmer, E., and Bally-Cuif, L. (2011). Expression of hairy/enhancer of split genes in neural progenitors and neurogenesis domains of the adult zebrafish brain. J Comp Neurol 519, 1748-69.

Chen, H., Smith, A.D., and Chen, T. (2016). WALT: fast and accurate read mapping for bisulfite sequencing. Bioinformatics 32, 3507-3509.

Collin, J., Queen, R., Zerti, D., Dorgau, B., Hussain, R., Coxhead, J., Cockell, S., and Lako, M. (2019). Deconstructing Retinal Organoids: Single Cell RNA-Seq Reveals the Cellular Components of Human Pluripotent Stem Cell-Derived Retina. Stem Cells 37, 593-598.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet 48, 1193-203.
Coulombre, J.L., and Coulombre, A.J. (1965). Regeneration of neural retina from the pigmented epithelium in the chick embryo. Dev Biol 12, 79-92.
Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48.
Estaras, C., Hsu, H.T., Huang, L., and Jones, K.A. (2017). YAP repression of the WNT3 gene controls hESC differentiation along the cardiac mesoderm lineage. Genes Dev 31, 2250-2263.
Fitzpatrick, D.R., and van Heyningen, V. (2005). Developmental eye disorders. Curr Opin Genet Dev 15, 348-53.
Fornes, O., Castro-Mondragon, J.A., Khan, A., van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Correard, S., Gheorghe, M., Baranasic, D., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res 48, D87-D92.
Fuhrmann, S. (2010). Eye morphogenesis and patterning of the optic vesicle. Curr Top Dev Biol 93, 61-84.
Fuhrmann, S., Levine, E.M., and Reh, T.A. (2000). Extraocular mesenchyme patterns the optic vesicle during early eye development in the embryonic chick. Development 127, 4599-609.
Garcia Delgado, A.B., de la Cerda, B., Alba Amador, J., Valdes Sanchez, M.L., Fernandez-Munoz, B., Relimpio Lopez, I., Rodriguez de la Rua, E., Diez Lloret, A., Calado, S.M., Sanchez Pernaute, R., et al. (2019). Subretinal Transplant of Induced Pluripotent Stem Cell-Derived Retinal Pigment Epithelium on Nanostructured Fibrin-Agarose. Tissue Eng Part A 25, 799-808.
Gehring, W.J. (2014). The evolution of vision. Wiley Interdiscip Rev Dev Biol 3, 1-40.
Gisselbrecht, S.S., Palagi, A., Kurland, J.V., Rogers, J.M., Ozadam, H., Zhan, Y., Dekker, J., and Bulyk, M.L. (2020). Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts. Mol Cell 77, 324-337 e8.
Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017-8.
Hatzfeld, M., Keil, R., and Magin, T.M. (2017). Desmosomes and Intermediate Filaments: Their Consequences for Tissue Mechanics. Cold Spring Harb Perspect Biol 9.
Heermann, S., Schutz, L., Lemke, S., Krieglstein, K., and Wittbrodt, J. (2015). Eye morphogenesis driven by epithelial flow into the optic cup facilitated by modulation of bone morphogenetic protein. Elife 4.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576-89.
Horsford, D.J., Nguyen, M.T., Sellar, G.C., Kothary, R., Arnheiter, H., and McInnes, R.R. (2005). Chx10 repression of Mitf is required for the maintenance of mammalian neuroretinal identity. Development 132, 177-87.
Hu, Y., Wang, X., Hu, B., Mao, Y., Chen, Y., Yan, L., Yong, J., Dong, J., Wei, Y., Wang, W., et al. (2019). Dissecting the transcriptome landscape of the human fetal neural retina and retinal pigment epithelium by single-cell RNA-seq analysis. PLoS Biol 17, e3000365.
Ivanovitch, K., Cavodeassi, F., and Wilson, S.W. (2013). Precocious acquisition of neuroepithelial character in the eye field underlies the onset of eye morphogenesis. Dev Cell 27, 293-305.
Kim, J.Y., Park, R., Lee, J.H., Shin, J., Nickas, J., Kim, S., and Cho, S.H. (2016). Yap is essential for retinal progenitor cell cycle progression and RPE cell fate acquisition in the developing mouse eye. Dev Biol 419, 336-347.

Kozmik, Z., Ruzickova, J., Jonasova, K., Matsumoto, Y., Vopalensky, P., Kozmikova, I., Strnad, H., Kawamura, S., Piatigorsky, J., Paces, V., et al. (2008). Assembly of the cnidarian camera-type eye from vertebrate-like components. Proc Natl Acad Sci U S A 105, 8989-93.
Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44, W90-7.
Kumar, L., and M, E.F. (2007). Mfuzz: a software package for soft clustering of microarray data. Bioinformation 2, 5-7.
Kwan, K.M., Otsuna, H., Kidokoro, H., Carney, K.R., Saijoh, Y., and Chien, C.B. (2012). A complex choreography of cell movements shapes the vertebrate eye. Development 139, 359-72.
Land, M.F., and Nilsson, D.-E. (2002). Animal eyes (Oxford ; New York: Oxford University Press)
Lane, B.M., and Lister, J.A. (2012). Otx but not Mitf transcription factors are required for zebrafish retinal pigment epithelium development. PLoS One 7, e49357.
Letelier, J., Bovolenta, P., and Martinez-Morales, J.R. (2017). The pigmented epithelium, a bright partner against photoreceptor degeneration. J Neurogenet 31, 203-215.
Li, Z., Joseph, N.M., and Easter, S.S.J. (2000). The morphogenesis of the zebrafish eye, including a fate map of the optic vesicle. Developmental Dynamics 218, 175-188.
Lian, I., Kim, J., Okazawa, H., Zhao, J., Zhao, B., Yu, J., Chinnaiyan, A., Israel, M.A., Goldstein, L.S., Abujarour, R., et al. (2010). The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation. Genes Dev 24, 1106-18.
Liu, I.S.C., Chen, J.D., Ploder, L., Vidgen, D., Vanderkooy, D., Kalnins, V.I., and Mcinnes, R.R. (1994). Developmental expression of a novel murine homeobox gene (Chx10): Evidence for roles in determination of the neuroretina and inner nuclear layer. Neuron 13, 377-393.
Loosli, F., Winkler, S., Burgtorf, C., Wurmbach, E., Ansorge, W., Henrich, T., Grabher, C., Arendt, D., Carl, M., Krone, A., et al. (2001). Medaka *eyeless* is the key factor linking retinal dertermination and eye growth. Development 128, 4035-4044.
Lowe, E.K., Cuomo, C., Voronov, D., and Arnone, M.I. (2019). Using ATAC-seq and RNA-seq to increase resolution in GRN connectivity. Methods Cell Biol 151, 115-126.
Magri, M.S., Jimenez-Gancedo, S., Bertrand, S., Madgwick, A., Escriva, H., Lemaire, P., and Gomez-Skarmeta, J.L. (2019). Assaying Chromatin Accessibility Using ATAC-Seq in Invertebrate Chordate Embryos. Front Cell Dev Biol 7, 372.
Manoli, M., and Driever, W. (2012). Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. Cold Spring Harb Protoc 2012.
Martinez-Morales, J.R. (2016). Vertebrate Eye Gene Regulatory Networks. In Organogenetic Gene Networks, J., C.-G.H. and P., B., ed. (Switzerland: Springer, Cham).
Martinez-Morales, J.R., Cavodeassi, F., and Bovolenta, P. (2017). Coordinated Morphogenetic Mechanisms Shape the Vertebrate Eye. Front Neurosci 11, 721.
Martinez-Morales, J.R., Dolez, V., Rodrigo, I., Zaccarini, R., Leconte, L., Bovolenta, P., and Saule, S. (2003). OTX2 activates the molecular network underlying retina pigment epithelium differentiation. J Biol Chem 278, 21721-31.
Martinez-Morales, J.R., Rodrigo, I., and Bovolenta, P. (2004). Eye development: a view from the retina pigmented epithelium. Bioessays 26, 766-77.
Martinez-Morales, J.R., Signore, M., Acampora, D., Simeone, A., and Bovolenta, P. (2001). Otx genes are required for tissue specification in the developing eye. Development 128, 2019-2030.
Masai, I., Stemple, D.L., Okamoto, H., and Wilson, S.W. (2000). Midline signals regulate retinal neurogenesis in zebrafish. Neuron 27, 251-263.
Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet 7, 29-59.
Mathers, P.H., Grinberg, A., Mahon, K.A., and Jamrich, M. (1997). The Rx homeobox gene is essential for vertebrate eye development. Nature 387, 603-607.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28, 495-501.
Miesfeld, J.B., Gestri, G., Clark, B.S., Flinn, M.A., Poole, R.J., Bader, J.R., Besharse, J.C., Wilson, S.W., and Link, B.A. (2015). Yap and Taz regulate retinal pigment epithelial cell fate. Development 142, 3021-32.
Mochii, M., Ono, T., Matsubara, Y., and Eguchi, G. (1998). Spontaneous transdifferentiation of quail pigmented epithelial cell is accompanied by a mutation in the Mitf gene. Dev Biol 196, 145-59.
Moreno-Marmol, T., Cavodeassi, F., and Bovolenta, P. (2018). Setting Eyes on the Retinal Pigment Epithelium. Front Cell Dev Biol 6, 145.
Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods 12, 982-8.
Nguyen, M.T., and Arnheiter, H. (2000). Signaling and transcriptional regulation in early mammalian eye development: a link between FGF and MITF. Development 127, 3581-3591.
Nicolas-Perez, M., Kuchling, F., Letelier, J., Polvillo, R., Wittbrodt, J., and Martinez-Morales, J.R. (2016). Analysis of cellular behavior and cytoskeletal dynamics reveal a constriction mechanism driving optic cup morphogenesis. Elife 5.
Owaribe, K., Kartenbeck, J., Rungger-Brandle, E., and Franke, W.W. (1988). Cytoskeletons of retinal pigment epithelial cells: interspecies differences of expression patterns indicate independence of cell function from the specific complement of cytoskeletal proteins. Cell Tissue Res 254, 301-15.
Picker, A., Cavodeassi, F., Machate, A., Bernauer, S., Hans, S., Abe, G., Kawakami, K., Wilson, S.W., and Brand, M. (2009). Dynamic coupling of pattern formation and morphogenesis in the developing vertebrate retina. PLoS Biol 7, e1000214.
Pittack, C., Jones, M., and Reh, T.A. (1991). Basic fibroblast growth factor induces retinal pigment epithelium to generate neural retina in vitro. Development 113, 577-88.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-2.
Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42, W187-91.
Raviv, S., Bharti, K., Rencus-Lazar, S., Cohen-Tayar, Y., Schyr, R., Evantal, N., Meshorer, E., Zilberberg, A., Idelson, M., Reubinoff, B., et al. (2014). PAX6 regulates melanogenesis in the retinal pigmented epithelium through feed-forward regulatory interactions with MITF. PLoS Genet 10, e1004360.
Reinhardt, R., Centanin, L., Tavhelidse, T., Inoue, D., Wittbrodt, B., Concordet, J.P., Martinez-Morales, J.R., and Wittbrodt, J. (2015). Sox2, Tlx, Gli3, and Her9 converge on Rx2 to define retinal stem cells in vivo. EMBO J 34, 1572-88.
Rembold, M., Loosli, F., Adams, R.J., and Wittbrodt, J. (2006). Individual cell migration serves as the driving force for optic vesicle evagination. Science 313, 1130-4.
Rowan, S., Chen, C.M., Young, T.L., Fisher, D.E., and Cepko, C.L. (2004). Transdifferentiation of the retina into pigmented cells in ocular retardation mice defines a new function of the homeodomain gene Chx10. Development 131, 5139-52.
Roy, A., de Melo, J., Chaturvedi, D., Thein, T., Cabrera-Socorro, A., Houart, C., Meyer, G., Blackshaw, S., and Tole, S. (2013). LHX2 is necessary for the maintenance of optic identity and for the progression of optic morphogenesis. J Neurosci 33, 6877-84.
Seth, A., Culverwell, J., Walkowicz, M., Toro, S., Rick, J.M., Neuhauss, S.C., Varga, Z.M., and Karlstrom, R.O. (2006). belladonna/(Ihx2) is required for neural patterning and midline axon guidance in the zebrafish forebrain. Development 133, 725-35.

Shankaran, S.S., Dahlem, T.J., Bisgrove, B.W., Yost, H.J., and Tristani-Firouzi, M. (2017). CRISPR/Cas9-Directed Gene Editing for the Generation of Loss-of-Function Mutants in High-Throughput Zebrafish F0 Screens. Curr Protoc Mol Biol 119, 31 9 1-31 9 22.
Sidhaye, J., and Norden, C. (2017). Concerted action of neuroepithelial basal shrinkage and active epithelial migration ensures efficient optic cup morphogenesis. Elife 6.
Stampfel, G., Kazmar, T., Frank, O., Wienerroither, S., Reiter, F., and Stark, A. (2015). Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature 528, 147-51.
Strauss, O. (2005). The retinal pigment epithelium in visual function. Physiol Rev 85, 845-81.
Tang, X., Gao, J., Jia, X., Zhao, W., Zhang, Y., Pan, W., and He, J. (2017). Bipotent progenitors as embryonic origin of retinal stem cells. J Cell Biol 216, 1833-1847.
Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032-4.
Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 13, 2129-41.
Vazquez-Marin, J., Gutierrez-Triana, J.A., Almuedo-Castillo, M., Buono, L., Gomez-Skarmeta, J.L., Mateo, J.L., Wittbrodt, J., and Martinez-Morales, J.R. (2019). yap1b, a divergent Yap/Taz family member, cooperates with yap1 in survival and morphogenesis via common transcriptional targets. Development 146.
Vejnar, C.E., Moreno-Mateos, M.A., Cifuentes, D., Bazzini, A.A., and Giraldez, A.J. (2016). Optimization Strategies for the CRISPR-Cas9 Genome-Editing System. Cold Spring Harb Protoc 2016.
Vopalensky, P., and Kozmik, Z. (2009). Eye evolution: common use and independent recruitment of genetic components. Philos Trans R Soc Lond B Biol Sci 364, 2819-32.
Waddington, C.H. (1966). Principles of development and differentiation (New York: New York, Macmillan)
Wang, J., Zibetti, C., Shang, P., Sripathi, S.R., Zhang, P., Cano, M., Hoang, T., Xia, S., Ji, H., Merbs, S.L., et al. (2018). ATAC-Seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. Nat Commun 9, 1364.
Wang, Z., Yasugi, S., and Ishii, Y. (2016). Chx10 functions as a regulator of molecular pathways controlling the regional identity in the primordial retina. Dev Biol 413, 104-11.
Westenskow, P., Piccolo, S., and Fuhrmann, S. (2009). Beta-catenin controls differentiation of the retinal pigment epithelium in the mouse optic cup by regulating Mitf and Otx2 expression. Development 136, 2505-10.
Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J.G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., et al. (2012). Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell 151, 1200-13.
Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., et al. (2019). The cis-Regulatory Atlas of the Mouse Immune System. Cell 176, 897-912 e20.
Zanconato, F., Forcato, M., Battilana, G., Azzolin, L., Quaranta, E., Bodega, B., Rosato, A., Bicciato, S., Cordenonsi, M., and Piccolo, S. (2015). Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. Nat Cell Biol 17, 1218-27.
Zou, C., and Levine, E.M. (2012). Vsx2 controls eye organogenesis and retinal progenitor identity via homeodomain and non-homeodomain residues required for high affinity DNA binding. PLoS Genet 8, e1002924.
Zuber, M.E., Gestri, G., Viczian, A.S., Barsacchi, G., and Harris, W.A. (2003). Specification of the vertebrate eye by a network of eye field transcription factors. Development 130, 5155-67.

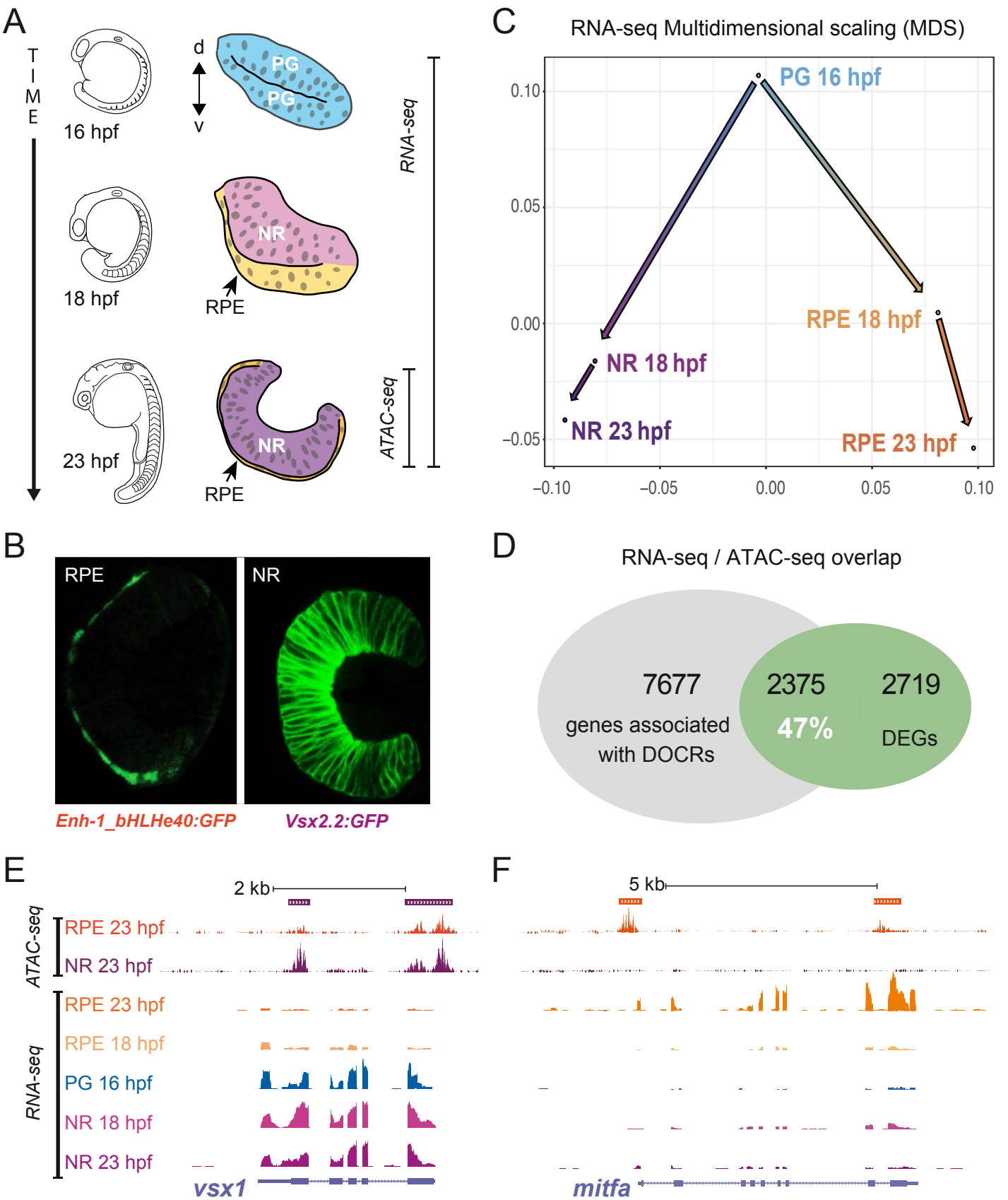
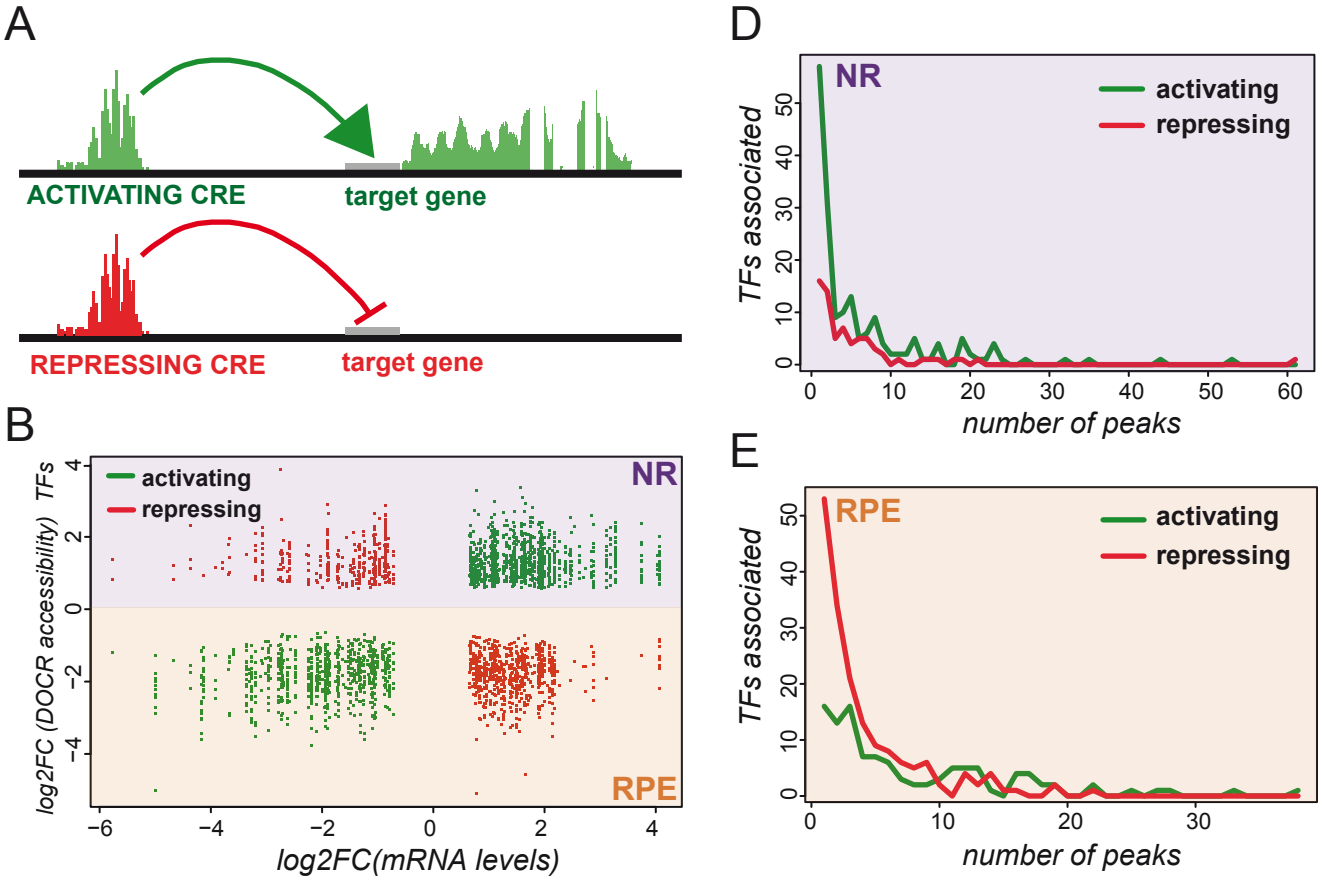


Figure 1: Experimental setup and raw data (A) Schematic representation of optic vesicle morphogenesis from undifferentiated retinal progenitors cells (PG, blue), at 16 hpf, to segregation and differentiation of the neural retina (NR, purple) and pigmented epithelium (RPE, orange) domains at 18 and 23 hpf. NR and RPE populations isolated by flow cytometry were analysed by RNA-seq and ATAC-seq at the stages indicated. (B) Stage 23 hpf confocal sections from the zebrafish transgenic lines used to mark and isolate the RPE (E1_bHLHe40:GFP) and NR (vsx2.2:GFP) populations. (C) Multidimensional scaling analysis of the RNA-seq data showing a progressive transcriptomic divergence between the NR and RPE domains. (D) Percentage of overlapping between differentially expressed genes (DEGs) and genes associated with differentially opened chromatin regions (DOCRs). (E-F) Overview of ATAC-seq and RNA-seq tracks (UCSC browser) for representative NR (vsx1; E) and RPE (mitfa; F) specific genes. Solid bars on the top indicate DOCRs. If purple, the DOCR is more accessible in the NR. On the contrary, if orange, the DOCR is more accessible in the RPE. In the depicted case all the DOCRs are accompanied by an increased transcription of the associated gene in the corresponding tissue.



C

	all genes				TFs		
	peaks	genes	peaks/gene		peaks	genes	peaks/gene
NR activating CREs	3434	754	4.55		1174	182	6.45
NR repressing CREs	1999	552	3.62		388	69	5.62
RPE activating CREs	5984	1119	5.35		869	110	7.90
RPE repressing CREs	3358	883	3.80		706	173	4.08

Figure 2: CRE configuration in the NR and RPE domains. (A) Schematic representation of the functional relationship between DOCRs, either activating (green) or repressing (red), and their associated DEGs. (B) The graph illustrates the correlation between the levels of differentially expressed TFs (log2FC) and the accessibility of their associated DOCRs (log2). (C) The table summarizes the number of activating or repressing CREs associated with either to all the DEGs or only with the differentially expressed TFs. (D-E) Histograms correlating the number of TFs associated with activating or repressing CREs to the number of peaks per gene, for both the NR (D) and RPE (E) domains.

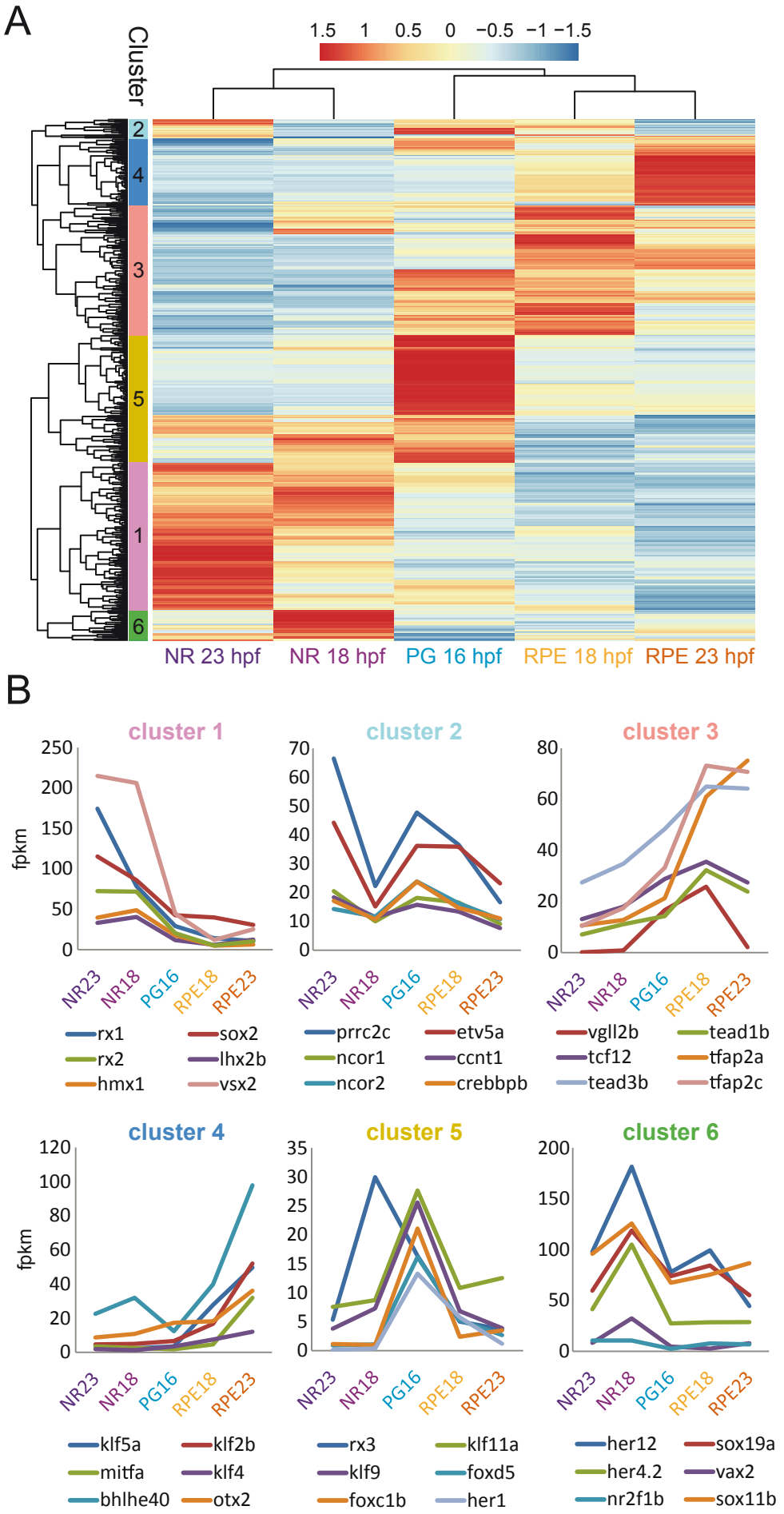


Figure 3: Hierarchical clustering of TFs gene expression variations during optic cup development. (A) Hierarchical clustering output shows TFs expression trends in the distinct domains and stages. Gene expression values, normalized by row, are indicated with a red to blue graded colour. Note that most clusters comprise a particular domain and developmental stage. (B) Relative transcript level changes for significant TFs (indicated by dotted lines) within each cluster.

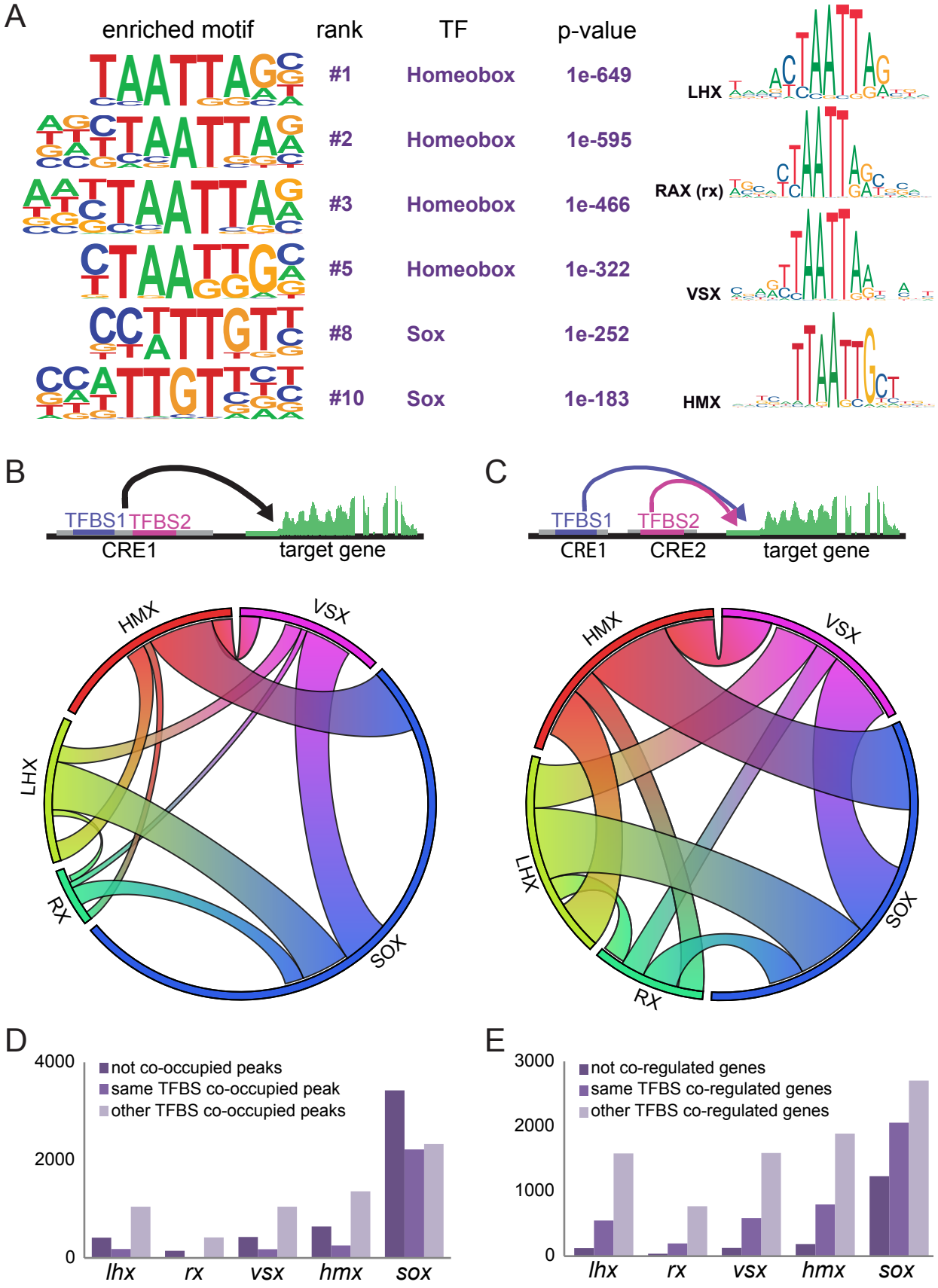


Figure 4: NR motif enrichment analysis. (A) Representative TF binding motifs enriched in NR DOCRs as identified by HOMER. The binding motif similarity among neural retina TFs of the homeobox family is indicated (http://jaspar.genereg.net). (B) Circoplot illustrating the co-occupancy rate of TFBS in the same DOCRs for the main TFs identified in the motif enrichment analysis (C) Circoplot illustrating the degree of co-regulation between TFs regulating the same gene through different DOCRs. (D) Number of CREs containing the main TFs identified in the motif enrichment analysis, classified according to their co-occupancy (E) Number of genes associated to CREs containing the main TFs identified in the motif enrichment analysis, classified according to their co-regulation.

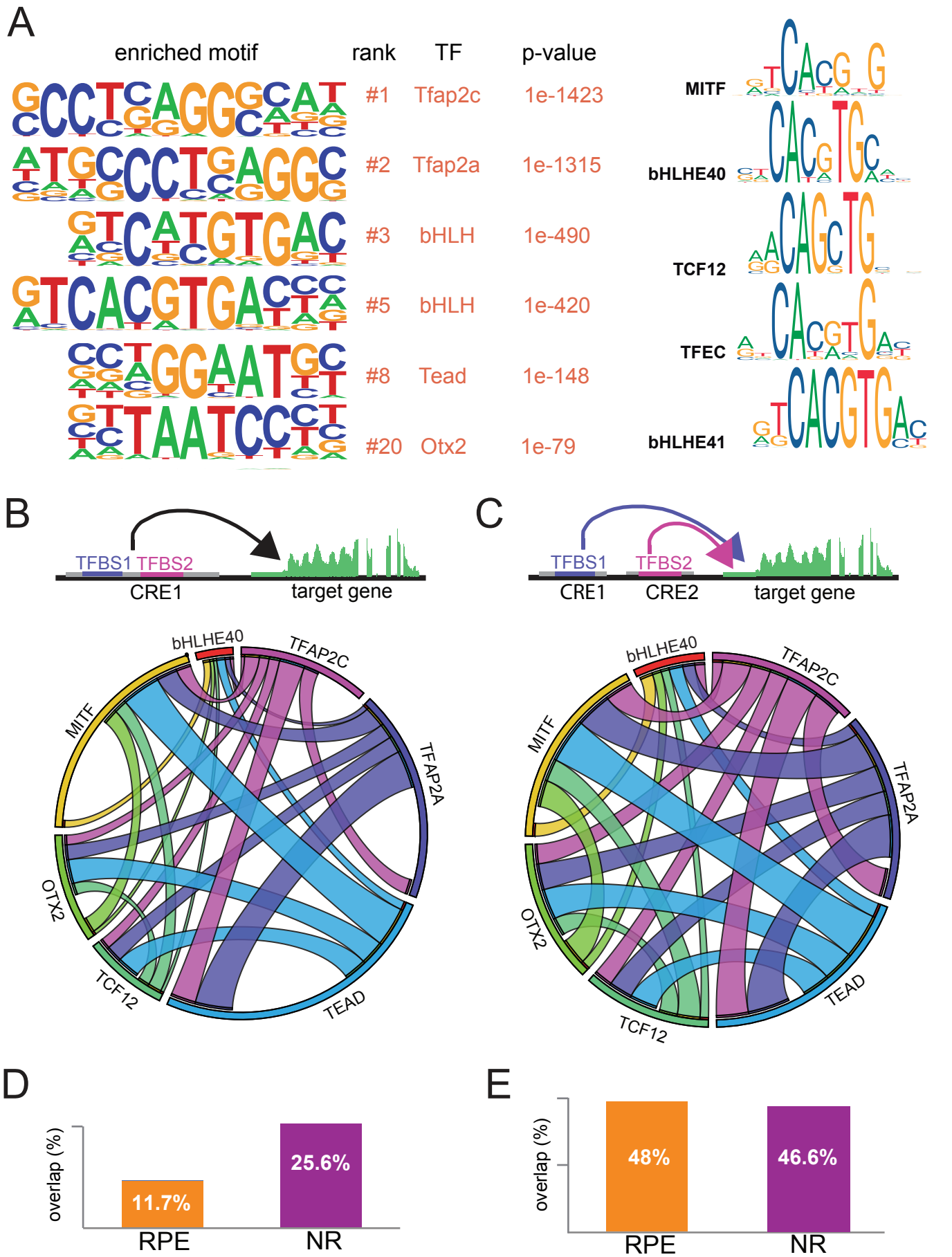


Figure 5: RPE motif enrichment analysis. (A) Representative TF binding motifs enriched in RPE DOCRs as identified by HOMER. Analysis of the binding motif similarity among TFs of the bHLH family (http://jaspar.genereg.net). (B) Circoplot illustrating the co-occupancy rate of TFBS in the same DOCRs for the main TFs identified in the motif enrichment analysis (C) Circoplot illustrating the degree of co-regulation between TFs regulating the same gene through different DOCRs. (D) Average percentage of co-occupancy in the same DOCR for two different TF in the RPE and NR. (E) Average percentage of two different TF regulating the same gene through different DOCRs in the RPE and NR.

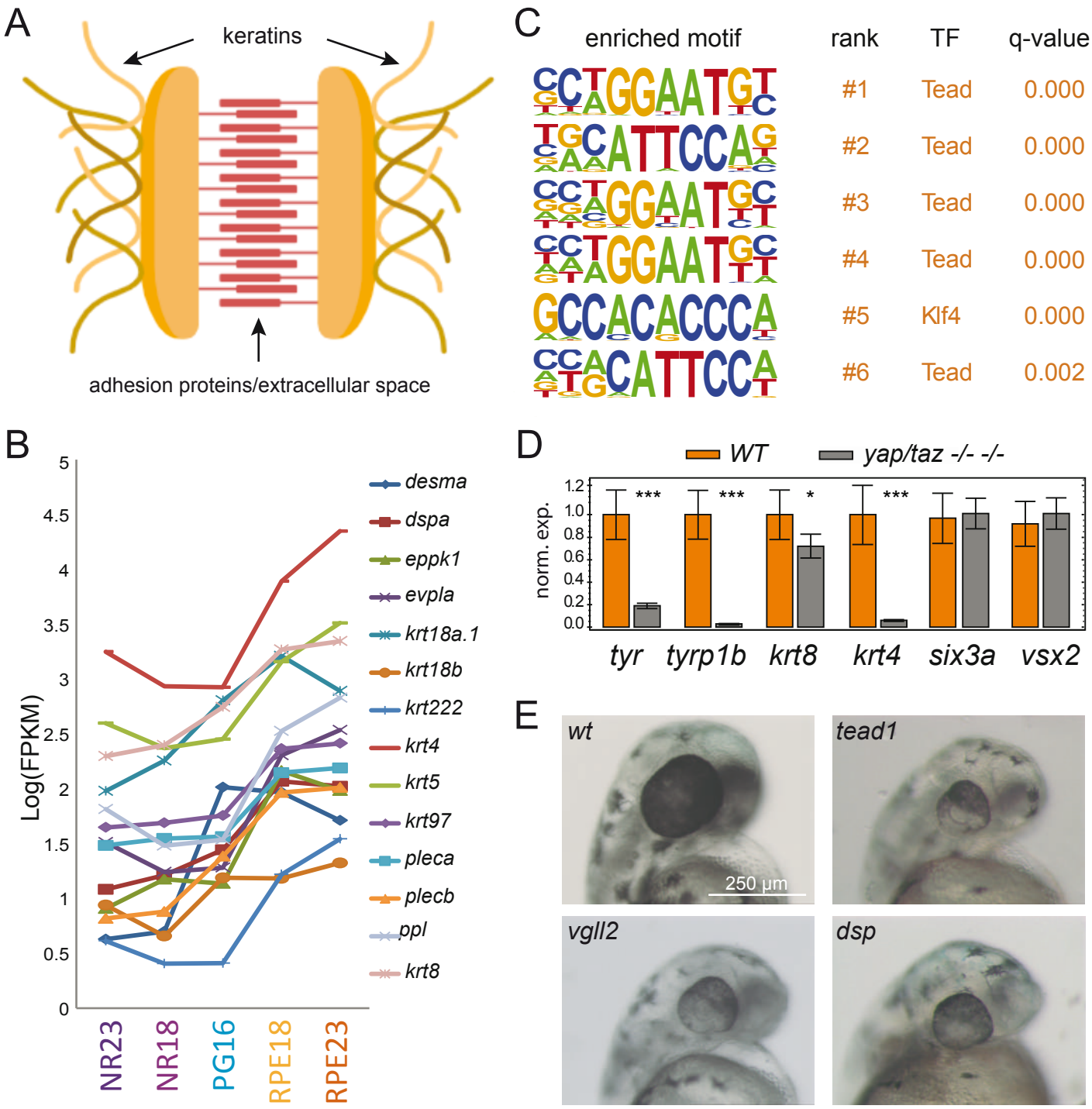


Figure 6: Regulation of desmosomal components during RPE specification. (A) Schematic representation of a desmosome junction. (B) Intermediate filament and desmosomal genes expression variations during optic cup development. Expression values are reported as log(FPKM) (C) Motif enrichment analysis of the DOCRs associated with genes encoding intermediate filament or desmosome components. (D) mRNA levels of keratin genes as well as RPE and NR markers as determined by RT-qPCR in wild type and yap -/- taz -/- double mutant zebrafish samples (dissected heads) at 18 hpf. Significant differences are indicated (n=3; T-test; ***= p<0.001). (E) Representative stereo microscope images of zebrafish embryo heads at 48 hpf. Magnification Bar = 0.5 mm): wild type and embryos injected with Cas9 (300 ng/µl) together with the following sgRNAs (80 ng/µl) combinations: vgll2a and vgll2b (vgll2); tead1a and tead1b (tead1) and dspa and dspb (dsp). Note the reduced eye size and RPE hypopigmentation in the crispants.

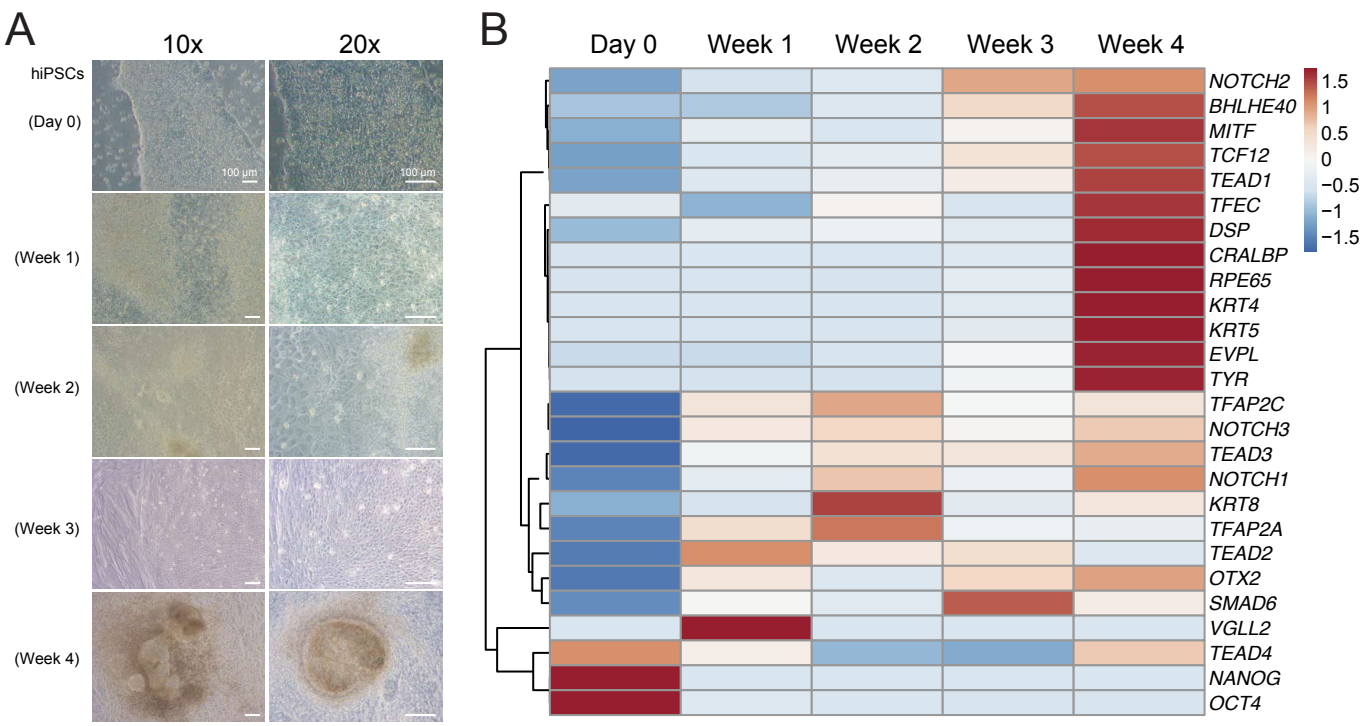
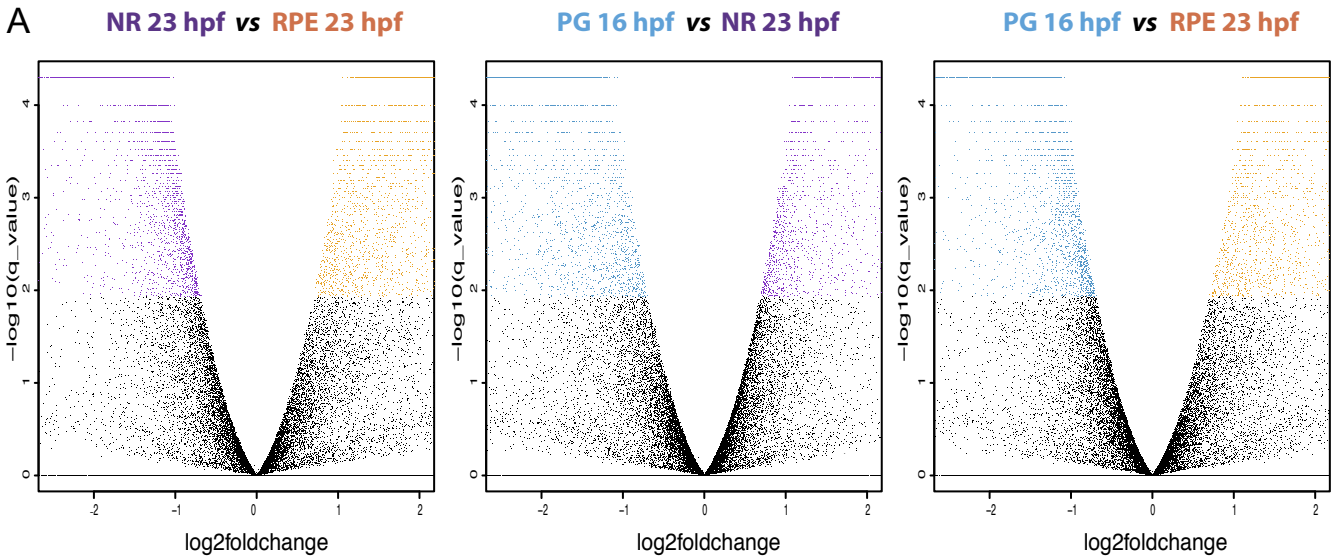


Figure 7: Gene expression during hiPSCs-to-RPE differentiation. (A) Bright field microscopy images (10-x and 20x) of hiPSCs before (Day 0) and during their differentiation towards RPE. Note the progressive acquisition of epithelial morphology and pigmentation. (B) Hierarchical clustered heatmap showing gene expression level variations, as determined by RT-qPCR during the differentiation towards RPE from hiPSCs. Note the conservation of the two clusters (early and late transcriptional waves) identified in zebrafish.

Figure S1

A

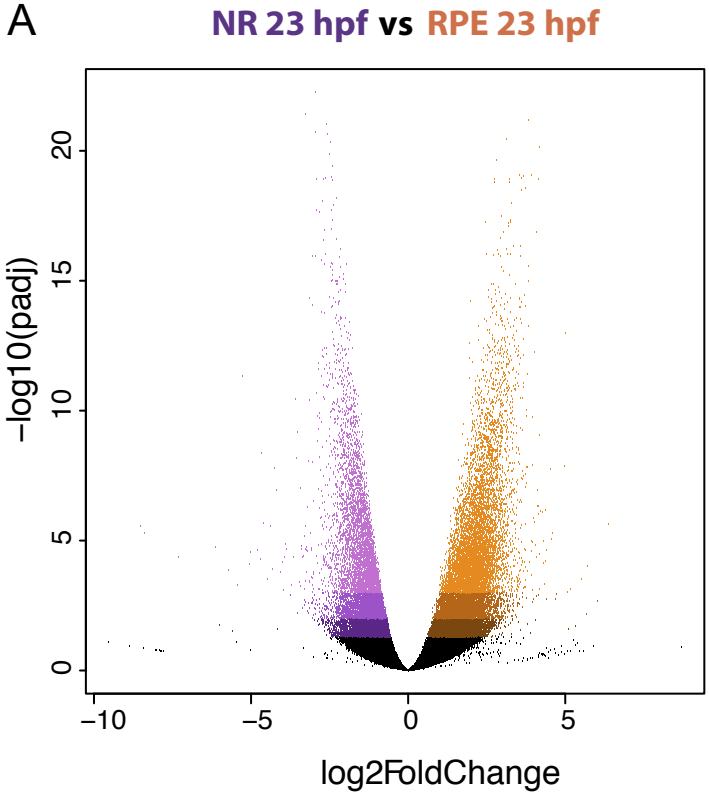


B

DOWN \ UP	PG 16 hpf	NR 18 hpf	RPE 18 hpf	NR 23 hpf	RPE 23 hpf
PG 16 hpf	-	650	1154	977	1518
NR 18 hpf	1474	-	1974	498	2105
RPE 18 hpf	1538	1808	-	2017	826
NR 23 hpf	1690	470	2105	-	2383
RPE 23 hpf	1849	1867	610	2158	-

Figure S1: Eye domain transcriptome variations during optic cup morphogenesis. (A) Volcano plots illustrating the transcriptome variations during eye morphogenesis. Each dot corresponds to a gene. Black dots indicate not significant variations, whereas coloured dots point out significant expression variations among domains and developmental stages. (B) Table summarizing the number of DEGs (upregulated or downregulated) in each one of the conditions.

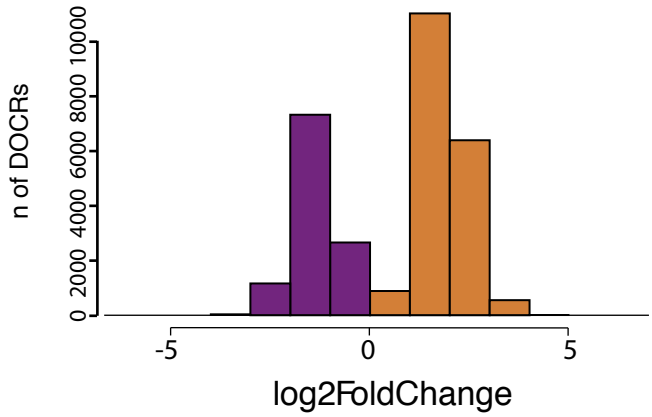
A



B

adj p-value	DOCRs	NR	RPE
<0.05	30172 **(12.6%)**	**11263**	**18909**
<0.01	19012 **(7.9%)**	**7062**	**11950**
<0.001	11387 **(4.8%)**	**4237**	**7150**

C



D

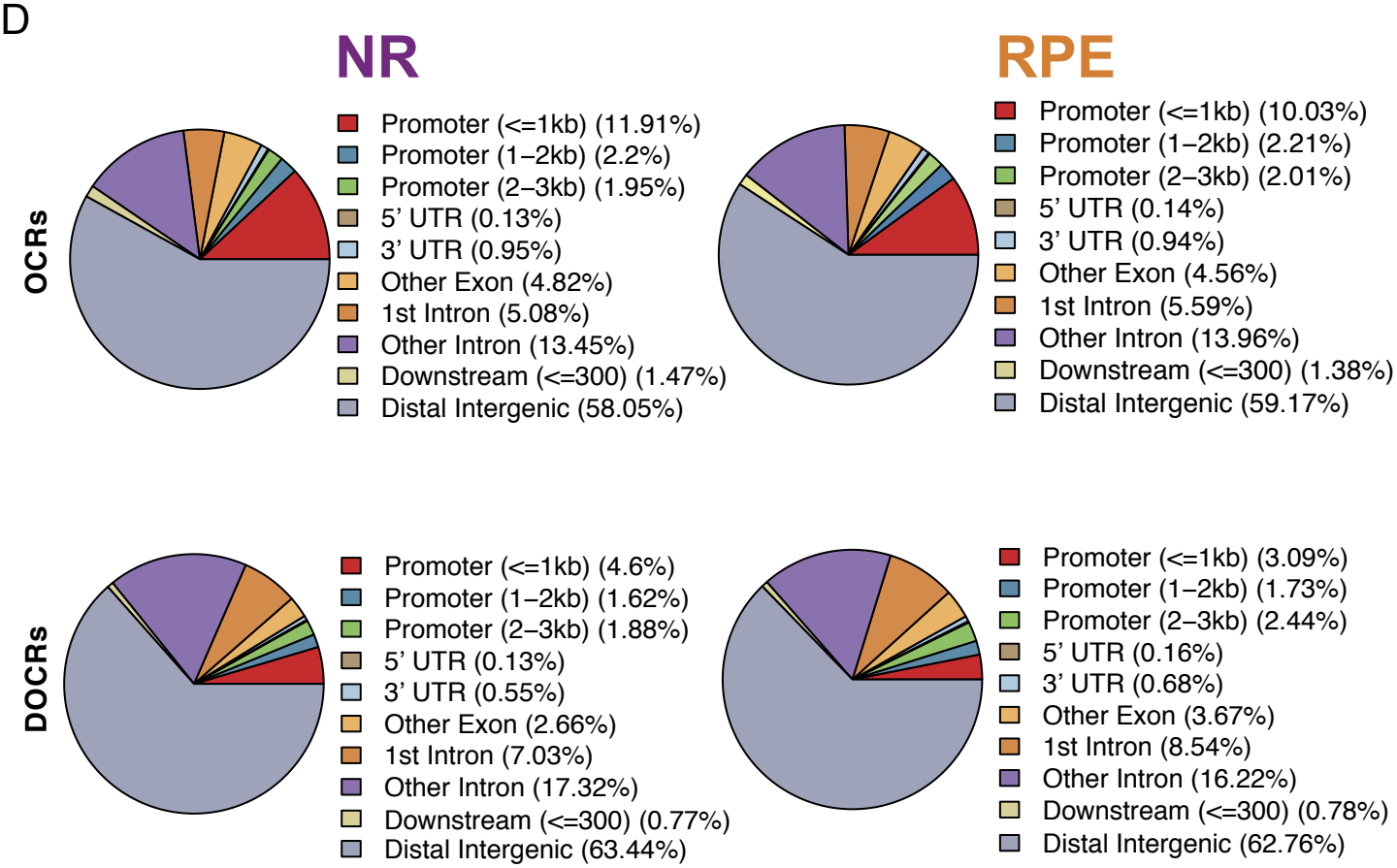


Figure S2: Analysis of differentially opened chromatin regions (DOCRs) during optic cup morphogenesis. (A) Volcano plot illustrating the chromatin accessibility changes during eye morphogenesis. Each dot corresponds to a peak (i.e. open chromatin region). Black dots indicate not significant variations; colour shades chromatin accessibility changes with different ranges of adjusted p-value (darker: p<0.05; medium= p<0.01, lighter= p<0.001). (B) Table including the number of peaks significantly more or less accessible between the two conditions. (C) Frequency histogram showing the distribution of the DOCRs in relation to their accessibility fold change. (D) Pie charts displaying the percentage of peaks falling in distinct regions of zebrafish genome. Top: genome distribution of the whole set of open chromatin regions (OCRs) identified by ATAC-seq in each condition. Bottom: genome distribution of the DOCRs.

Figure S3

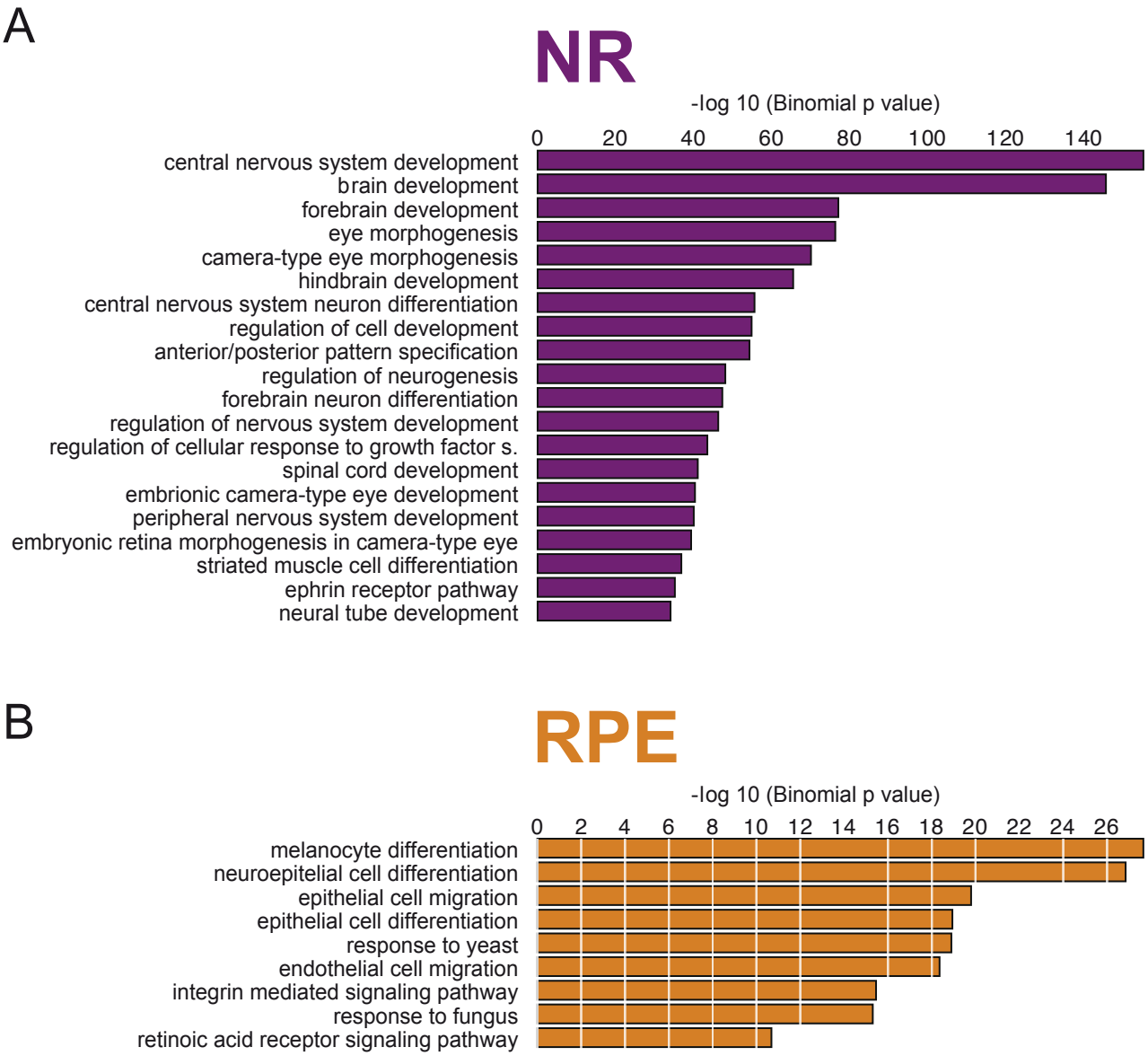


Figure S3: Gene ontology enrichment of the genes associated with DOCRs. Bar chart showing GO terms for biological processes enriched in the genes associated to DOCRs. Bar length is proportional to enrichment significance.

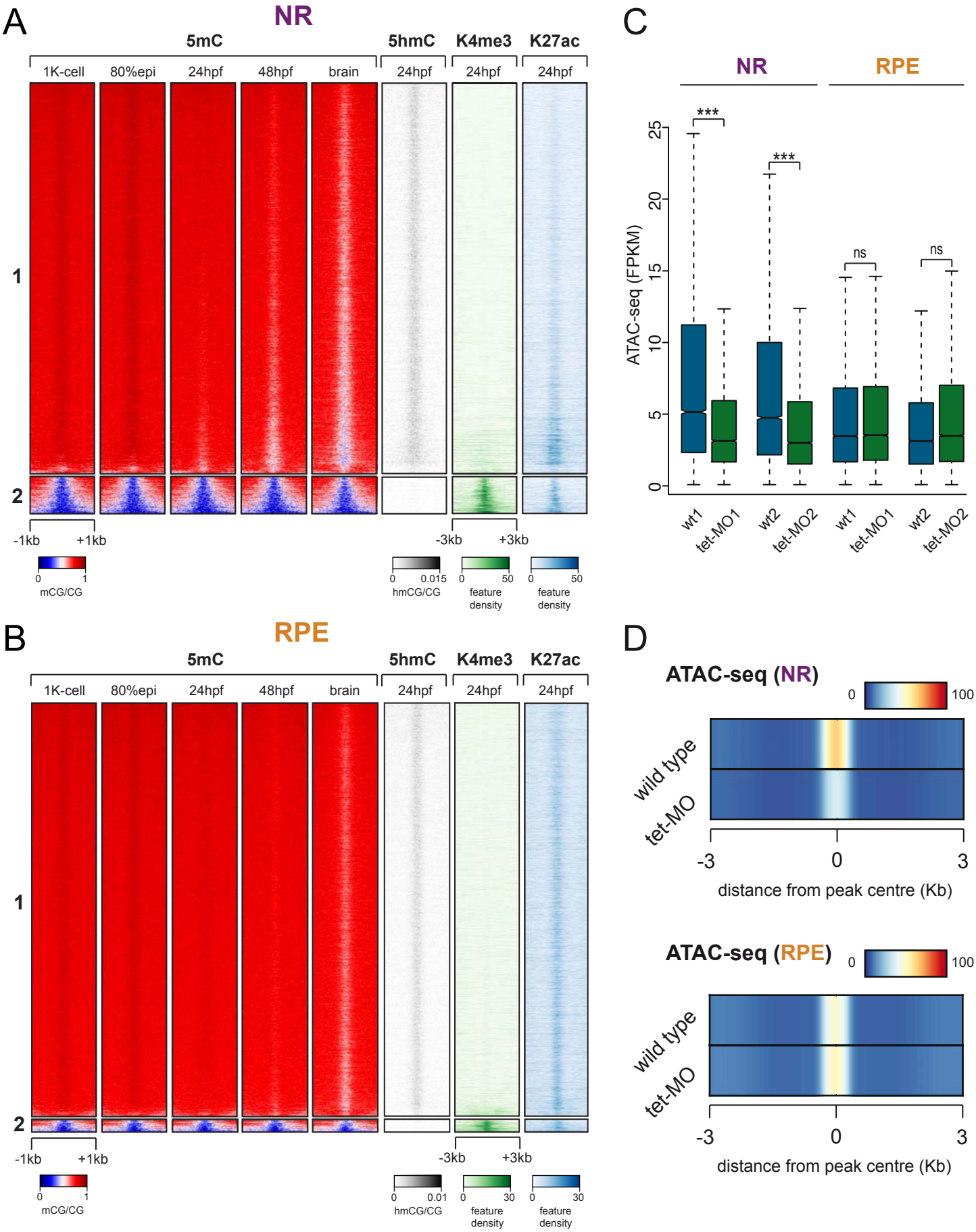


Figure S4. DNA methylation profiles of NR and RPE peaks. Clustering of NR (A) and RPE (B) ATAC-seq peaks according to 5mC pattern across developmental stages and in adult brain (left columns), as well as 5hmC, K4me3, and K27ac profiles at 24 hpf (right columns). Note that the signature of cluster #1 corresponds to active enhancers, whereas that of cluster #2 corresponds to hypomethylated promoter regions. (C) Chromatin accessibility quantification for NR and RPE peaks in wild type and triple tet (tet1/2/3) morphants. Analysis was performed in whole 24hpf embryos. Statistical significance was evaluated using a Mann-Whitney-Wilcoxon Test, $P < 2.2e-16$. (D) Chromatin accessibility levels for NR and RPE ATAC-seq peaks in wild type and and triple tet (tet1/2/3) zebrafish morphants.

Figure S5

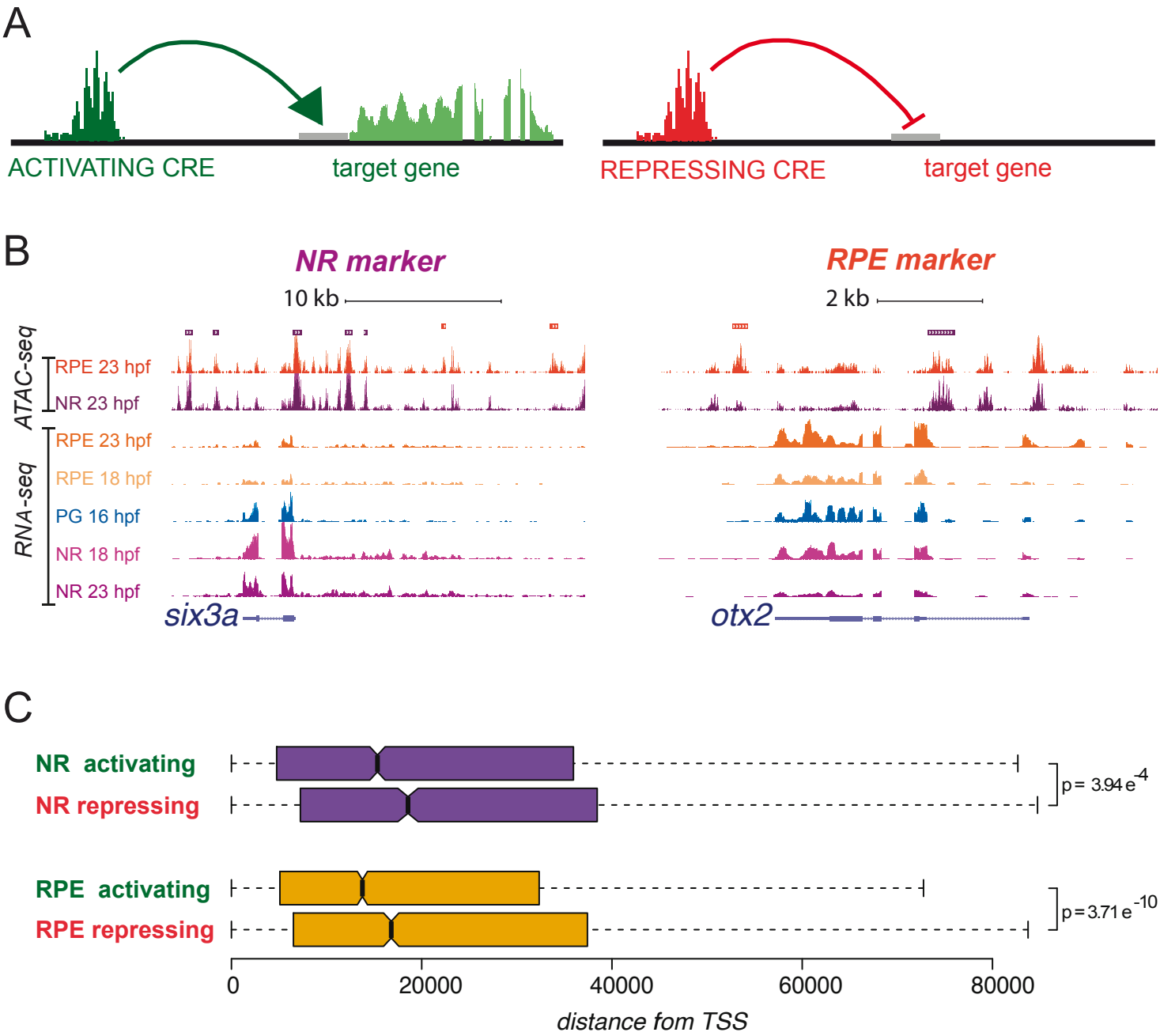


Figure S5: Activating and repressing CREs localization. (A) Schematic representation of the functional relations among DOCRs, either activating (green) or repressing (red), and their associated DEGs. (B) Illustrative examples of ATAC-seq and RNA-seq tracks (UCSC browser) for representative NR (six3; E) and RPE (otx2; F) genes. Bars on the top indicate DOCRs. If purple, the DOCR is more accessible in NR; if orange, the DOCR is more accessible in RPE. Some of the more accessible DOCRs are accompanied by a decreased transcription of the associated gene in the corresponding tissue. (C) Average distance to the TSS of all activating and repressing DOCRs for the NR and RPE domains (see Figure 2C). Note that activating CREs are significantly (T-test) closer to the TSS than repressing regions in both domains.

Figure S6

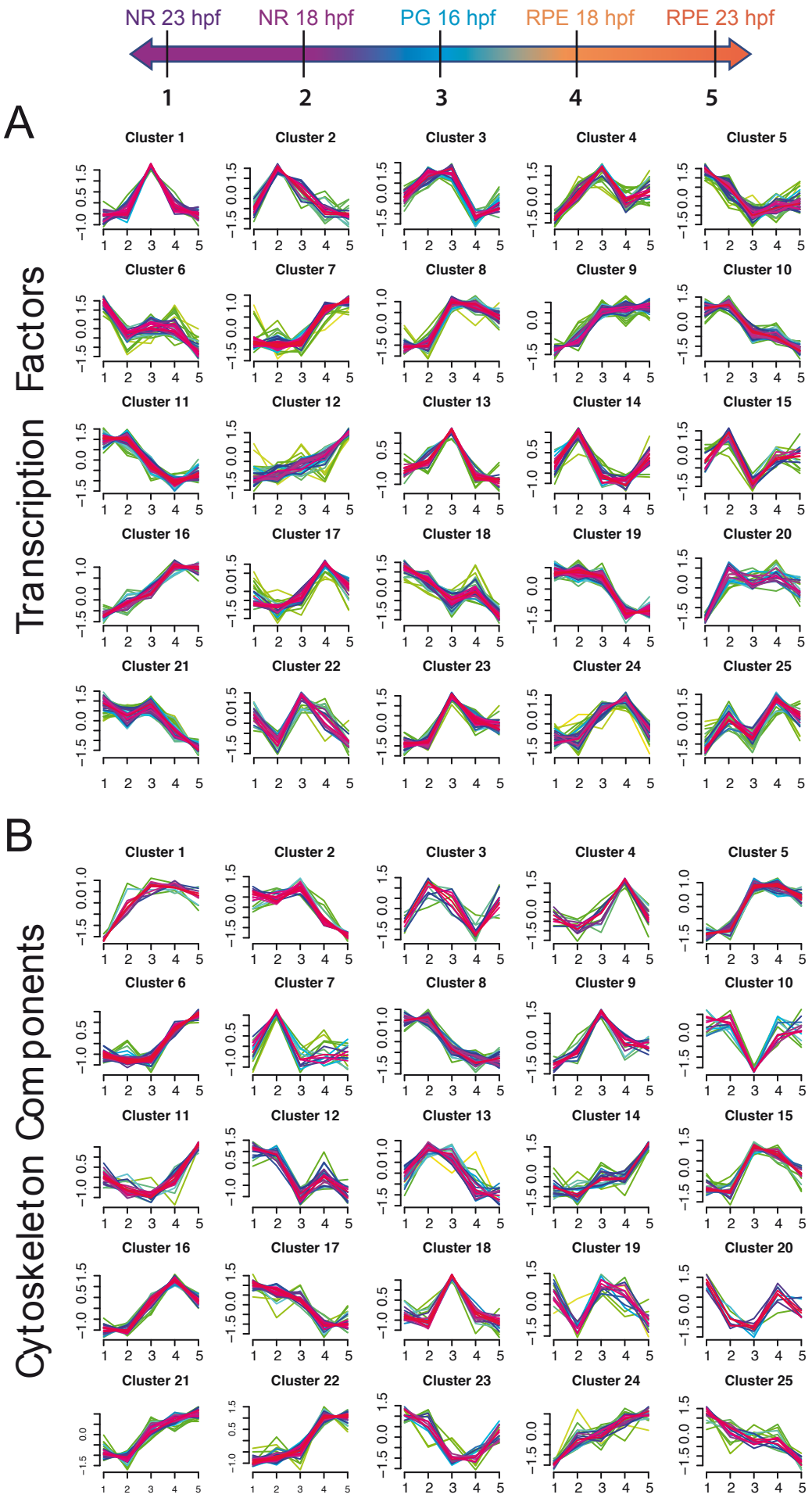


Figure S6: Partitioning clustering of gene expression variations during optic cup development. Partitioning clustering output (k= 25) showing the expression trends in the distinct domains and stages examined for two classes of genes: TFs (A) and cytoskeleton components (B). Stages and domains examined are linearly represented with a colour code and numbered from 1 (NR 23hpf) to 5 (RPE 23hpf), following the lineage trajectory.

Figure S7

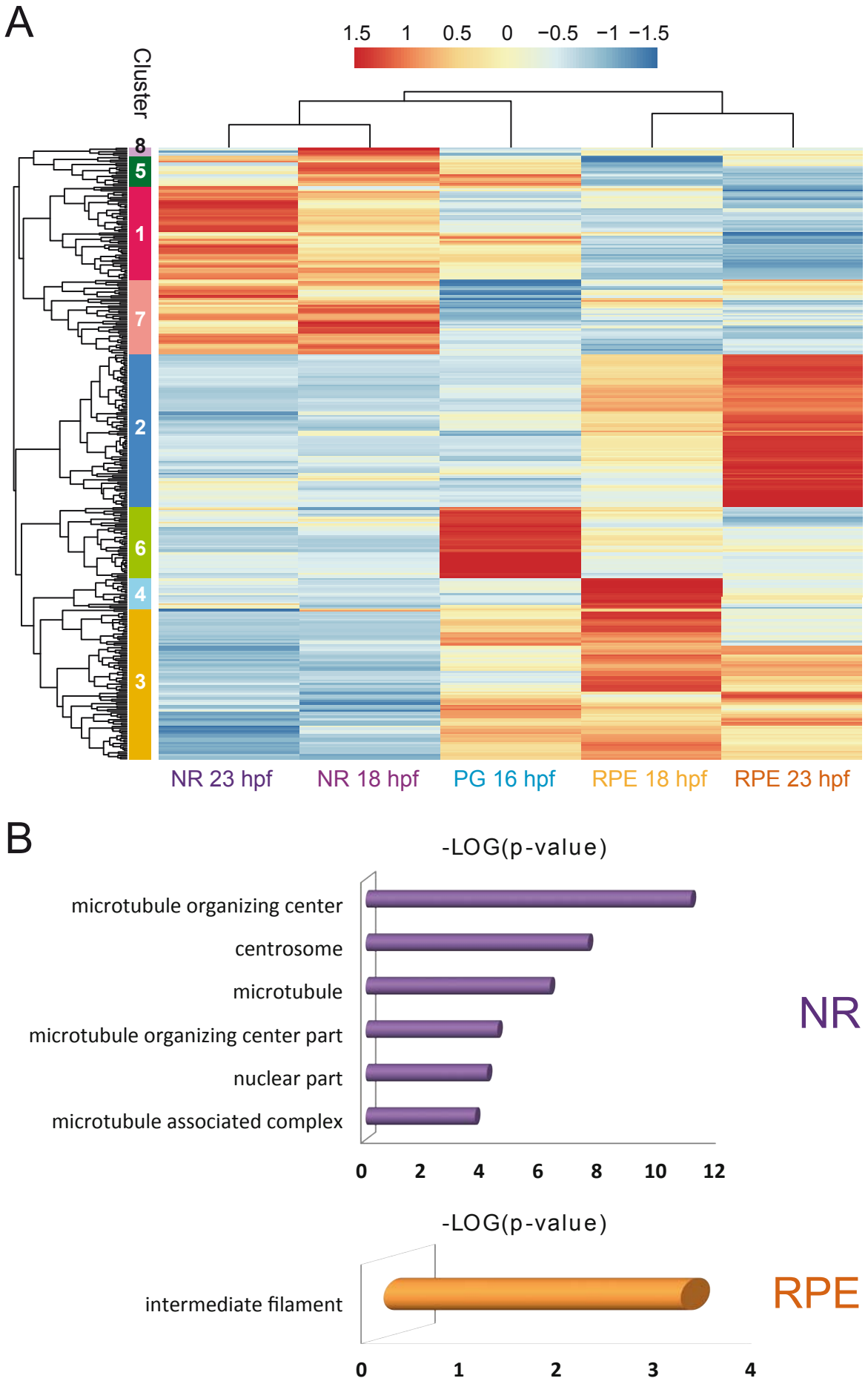


Figure S7: Hierarchical clustering of gene expression variations during optic cup development for cytoskeletal components. (A) Hierarchical clustering output showing the expression trends in distinct domains and stages. Gene expression values, normalized by row, are indicated with a red to blue graded colour. Note that all stages and domains are represented by at least one gene cluster (B) Cellular component GO enrichment for cytoskeletal genes belonging to NR or RPE clusters.

Figure S8

Figure S8: GO terms associated with RPE TFs. Hierarchical clustering of the main RPE TFs according to the enriched GO terms (Biological process) associated to their putative downstream genes (genes associated with RPE DOCRs that contain each TFBS).

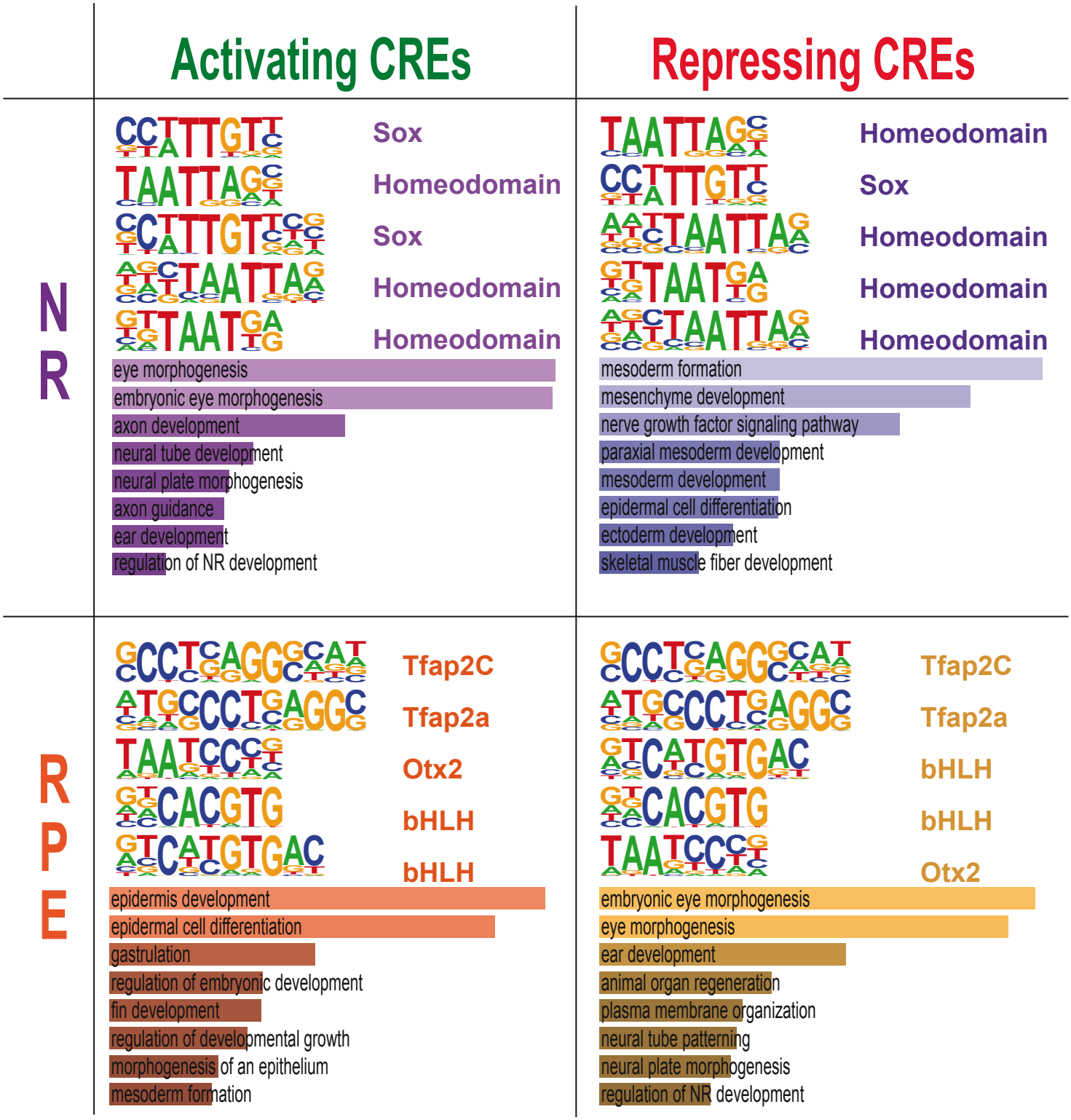
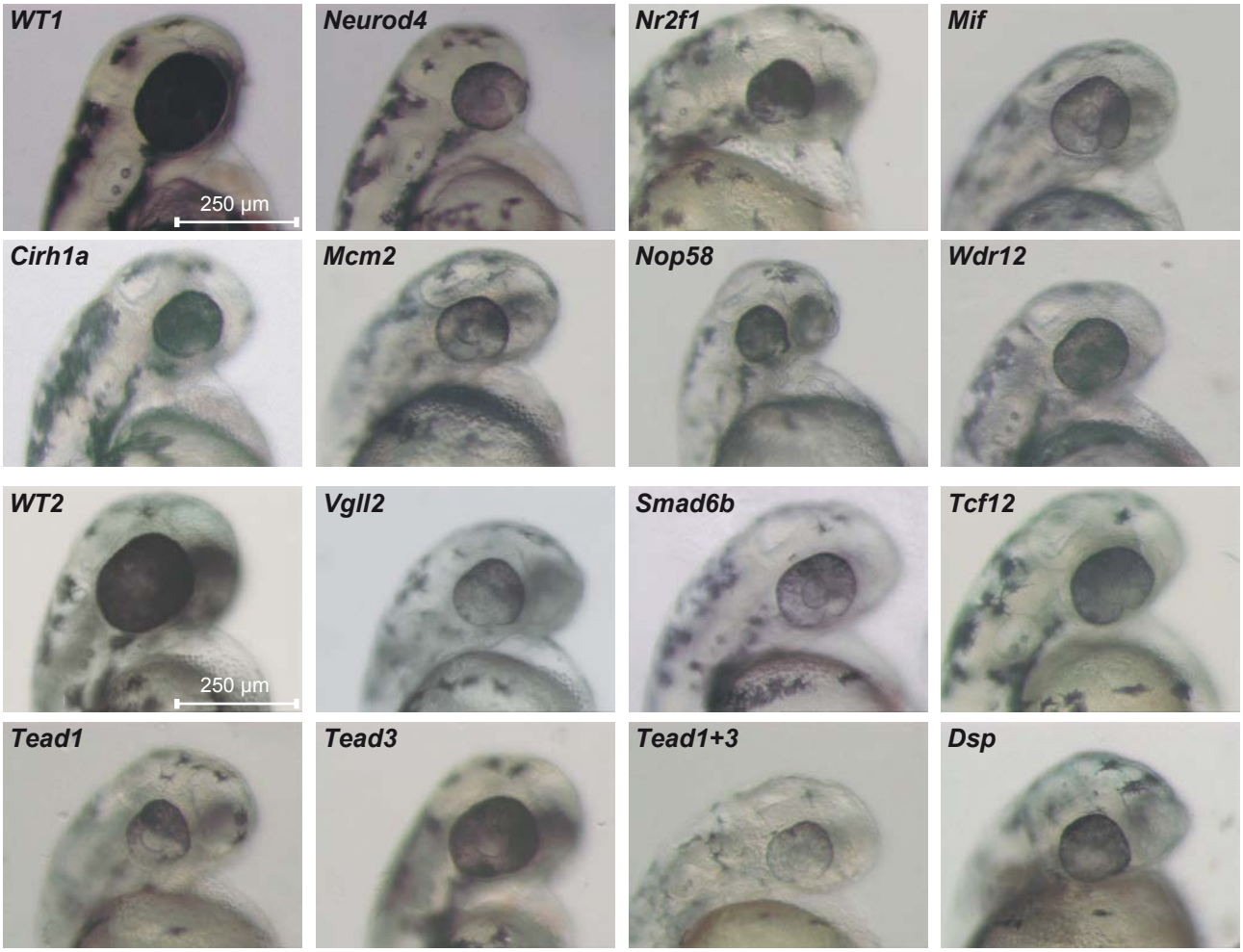


Figure S9: Motif enrichment analysis of activating and repressing CREs. Representative TF motifs enriched in activating (left column) and repressing (right column) CREs in both the NR (A) and RPE (B) domains. Analysis of GO terms enrichment for genes associated with each set of elements is indicated.

A



B

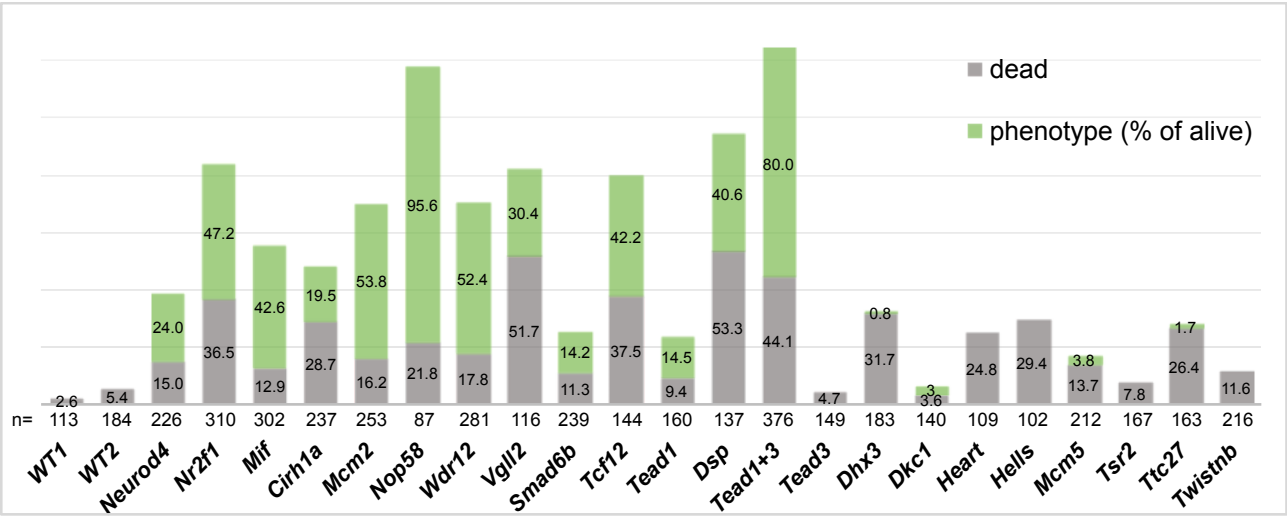


Figure S10: CRISPR CRISPR/Cas9 F0 screening. (A) Eye phenotypes resulting from the injection of Cas9 protein and sgRNAs for the candidate genes. (B) Percentage of embryos showing an impaired eye phenotype upon the injection of Cas9/sgRNAs complex (n= total number of injected embryos). The % of lethality upon injection is also shown. See also Supplementary data set 12.

Appendix Supplementary Tables

Table SI. Primers used for Real Time q-PCR of human samples.

gene	gene ID	Fw_primer (5'->3')	Rv_primer (5'->3')
OCT4	ENSG00000204531	CTTCAGGAGATATGCAAAGCAGA	TGATCTGCTGCAGTGTGGG
NANOG	ENSG00000111704	GGATCCAGCTTGTCCCCAAA	AGGAAGGAAGAGGAGAGACAGT
RPE65	ENSG00000116745	ACCACCTGTTTGATGGGCAA	AGTGCGGATGAACCTTCTGT
CRALBP	ENSG00000116745	GTCACAACTTGGCCCTGACT	GGTCCATGGTCCTTGGTTGT
TYR	ENSG00000077498	GATTCAGACCCAGACTCTTTTCA	ACGACACAGCAAGCTCACAA
HPRT1	ENSG00000165704	CCCTGGCGTCGTGATTAGTG	TCGAGCAAGACGTTCAGTCC
GAPDH	ENSG00000111640	AGGTCGGAGTCAACGGATTT	TGGAATTTGCCATGGGTGGA
MITF	ENSG00000187098	CCGGGCTCTGTTCTCACTTT	GGAACTGCTGCTCTTCAGCG
OTX2	ENSG00000165588	CCTCACTCGCCACATCTACT	AGTGGAACTTACAGCCTCATGG
BHLHE40	ENSG00000134107	ATTAACGAGTGCATCGCCCA	AGCTCACCAGCTTGTAAACCA
TFEC	ENSG00000105967	GATAAAATCCACTCATTGCTGGTCC	GGGCTTTCTGTAGCTGAGGC
TFAP2A	ENSG00000137203	GAGAGTAGCTCCACTTGGGTG	CCGTCGTGACGGTCCTCG
TFAP2C	ENSG00000087510	GAAGAGGACTGCGAGGATCG	GCTGATATTCGGCGACTCCA
TEAD1	ENSG00000187079	CCATTCCAGGGTTTGAGCCT	GCTTGGTTGTGCCAATGGAG
TEAD2	ENSG00000074219	TCGGAATGAACTGATCGCCC	CCTGGTCCTTCAACTTGGACT
TEAD3	ENSG00000007866	GACCGTACCATTGCCTCCTC	TTGCTGTACGTGTCAGGGTC
TEAD4	ENSG00000197905	GGCACCATTACCTCCAACGA	CAGCTCGTTCCGACCATACA
TCF12	ENSG00000140262	CCATGAAGGCTTGTCCCCAA	GGAGACTAGATTGACAGCCTGG
VGLL2	ENSG00000170162	GCTTTGCTCCGCCTGATGAC	ATAGGCTAGTTTCTGGTGGTAGG
SMAD6	ENSG00000137834	GGGCCCGAATCTCCGC	GGTCGTACACCGCATAGAGG
KRT5	ENSG00000186081	CGAGGAATGCAGACTCAGTG	GCTGCTGGAGTAGTAGCTTCC
KRT4	ENSG00000170477	TCCTTCATCGACAAGGTGCAG	GGGCTCAAGGTTTTTGCTGG
KRT8	ENSG00000170421	CAGCAAATGTTTGCGGAATGAA	AACCAGGCGGAGATCCCTTC
DSP	ENSG00000096696	AGGCTGGAGTACGATGACCT	TAGATGCCTCTAAAGCCTGC
EVPL	ENSG00000167880	CGACTTCCGACTGCTCCATCT	CCAAGTCCTCCAAGGGTGTG
NOTCH1	ENSG00000148400	CTGCCTCTTCGACGGCTTT	AAGTGGAAGGAGCTGTTGCG
NOTCH2	ENSG00000134250	CGAGTGTGTCCCAGGCTATC	CTTCACAGAGTAGGCCCCGA
NOTCH3	ENSG00000074181	GTCTTCCTGGGTTTGAGGGTC	GGGCACTGGCAGTTATAGGT

Table SII. Primers used for Real Time q-PCR of zebrafish samples.

gene	gene ID	Fw_primer (5'->3')	Rv_primer (5'->3')
tyr	ENSDARG00000039077	ACGGATACTTCATGGTGCCC	CGCTGACCTGGATCCTGTAAAT
tyrp1b	ENSDARG00000056151	GCCCGTCCAATGGTTCAAAG	GGAGCGCTGTAACCCTCAAT
krt4	ENSDARG00000017624	CTTCGTTGCGGCTCCTATCA	TCCAGGAAGCGCACTTTGTC
krt8	ENSDARG00000058358	TCCGCGCTCAGTATGAAGAC	AAGTTGGCTCGCTGTCCTTT
six3a	ENSDARG00000058008	AAAAACAGGCTCCAGCATCAA	AAGAATTGACGTGCCCGTGT
vsx2	ENSDARG00000005574	GGGATTAATTGGGCCTGGAGG	GCTGGCAGACTGGTTATGTTCC
eef1a1l1	ENSDARG00000020850	TCCACCGGTCACCTGATCTAC	CAACACCCAGGCGTACTTGA

Table SIII. sgRNAs used for CRISPR/Cas9 screen

gene	gene ID	sgRNA_seq (5'->3')
mphosh10	ENSDARG00000053912	GGTGGCTTTCGTGGACGAGGCGG
mphosh10	ENSDARG00000053912	GGATTTCGAGGAGGCAGGGGTGG
heatr1	ENSDARG00000099742	GAGGTGCTGGCTCTCCGTCATGG
heatr1	ENSDARG00000099742	GAGGGAGGGCCAATCAGCAAAGG
hells	ENSDARG00000057738	TGGGGCTGCTGTGCTGGCACAGG
hells	ENSDARG00000057738	AGACAGGTTATTCTGGAGGGGGG
nop58	ENSDARG00000104353	AGAGATCTCGATGGGCACAGAGG
nop58	ENSDARG00000104353	GGGCATCAGAAACCAGATGGAGG
mcm5	ENSDARG00000019507	GCGTAACCCTGCAGCCCCGGTGG
mcm5	ENSDARG00000019507	GTGGCGCAGACCAAAGCCAAAGG
dkc1	ENSDARG00000016484	GAGCTGCGACGAGTCCGTTCCGG
dkc1	ENSDARG00000016484	GGGCTGCCTGATCGTGTGTGTGG
cirh1a	ENSDARG00000017675	GGGCCAATCTGGGCCATAACAGG
cirh1a	ENSDARG00000017675	AGAGGGTACGGGACGTCCCGCGG
wdr12	ENSDARG00000003287	GGGGAAGGCTGTGATGACTGTGG
wdr12	ENSDARG00000003287	GAGATCTGCAACCTCGGAGGAGG
ttc27	ENSDARG00000007918	GGAAGTTGCTCTGTTGGCGGTGG
ttc27	ENSDARG00000007918	GTGGCTCCTCTGCTCTTCGGTGG
dhx33	ENSDARG00000051785	GAGGCGGGCATCGGCCGGCAGGG
dhx33	ENSDARG00000051785	GTGTTTGGAGATGTCCCGGCAGG
mcm2	ENSDARG00000102798	GGGCCACACGGTGCGCGAGTGGG
mcm2	ENSDARG00000102798	GAGCGACTGACACTCAGGACAGG
tsr2	ENSDARG00000005772	GTGTGAGCAGGGCAGATTGGCGG
tsr2	ENSDARG00000005772	TGGAGCGTTCAGTCAGCAGAAGG
mif	ENSDARG00000071336	GTGACAGTACATCGCCGTACAGG
mif	ENSDARG00000071336	GTGAGCGAGCAGAGCGCACACGG
fbl	ENSDARG00000053912	GGTGGCTTTCGTGGACGAGGCGG
fbl	ENSDARG00000053912	GGATTTCGAGGAGGCAGGGGTGG
tcf12	ENSDARG00000004714	AGCTTCCAGTGGCGTTCCCTCGG
tcf12	ENSDARG00000004714	GTGGGCGACACCGAGTGTGGCGG
smad6b	ENSDARG00000031763	TGTGCTGCAGGTCAGACCACCGG
smad6b	ENSDARG00000031763	GGGGAAAGTCTTGAGTATGGAGG
vgll2a	ENSDARG00000041706	AGGGGACATCAGTTCGGTGGTGG
vgll2a	ENSDARG00000041706	GTGTATGCGGCTGCAAAATACGG
vgll2b	ENSDARG00000053773	TGGAGCCAGGTAAGCTGATGAGG
vgll2b	ENSDARG00000053773	GAGGTCGGCTCGGGGGAGAGGGG
neurod4	ENSDARG00000003469	TGGCTTTGATTCGGCGGGCACGG
neurod4	ENSDARG00000003469	TGGTTGTGGGCCCAAGTTGGAGG
nr2f1a	ENSDARG00000052695	AGCGCATACTGGCCCGGGTTCGG
nr2f1a	ENSDARG00000052695	GCCGTCCCTGGTGTGGACGGAGG
nr2f1b	ENSDARG00000017168	TGCGGTGGTGCTGATCCACCGGG
nr2f1b	ENSDARG00000017168	TGGCTCGGGTTCGGCTGGTTCGG
tead1a	ENSDARG00000028159	TGTGTCGTTGAAGGATCATACGG
tead1a	ENSDARG00000028159	CGGCAGTGAAAGTGCCGGGGAGG
tead1b	ENSDARG00000059483	GACACCTGCGGGGTAGGGGAGGG
tead1b	ENSDARG00000059483	AGAGGCCGGTCTTACCCCTGCGG
tead3a	ENSDARG00000074321	GATGATCTTTCTGCGGCCACAGG
tead3a	ENSDARG00000074321	TGAAGGGTAGGGTACGGGCTCGG
tead3b	ENSDARG00000063649	TGAAGGTATGCGCTTTCCTGCGG
tead3b	ENSDARG00000063649	GATGGGGGTCGGCCAGAACTGGG
dspa	ENSDARG00000022309	TGGCACGTGACTGGACCTGGAGG
dspa	ENSDARG00000022309	AGGCTGATGCTCAGAGGGTTTGG
wu:fi04e1	ENSDARG00000076673	AGGATCTGAATATTCAGCGGCGG
wu:fi04e1	ENSDARG00000076673	GACGACGGCCTGGAGGAGTGCGG