Widespread translational control regulates retinal development in mouse

Kaining Chen[1,†], Congying Chen[1,†], Huihui Li[1], Jiaqi Yang[1], Mengqing Xiang[1], Hongwei Wang[1,*], Zhi Xie[1,*]

[1] State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China

† These authors contributed equally to this work.

* To whom correspondence should be addressed. Email: bioccwhw@126.com or xiezhi@gmail.com

Abstract

Retinal development is tightly regulated to ensure the generation of appropriate cell types and the assembly of functional neuronal circuitry. Despite remarkable advances have been made in understanding regulation of gene expression during retinal development, how translational regulation guides retinogenesis is less understood. Here, we conduct a comprehensive translatome and transcriptome survey to the mouse retinogenesis from the embryonic to the adult stages. We discover thousands of genes that have dynamic changes at the translational level and pervasive translational regulation in a developmental stage-specific manner with specific biological functions. We further identify genes whose translational efficiencies are frequently controlled by changing usage in upstream open reading frame during retinal development. These genes are enriched for biological functions highly important to neurons, such as neuron projection organization and microtubule-based protein transport. Surprisingly, we discover hundreds of previously undetected microproteins, translated from long non-coding RNA and circular RNAs. We validate their protein products *in vitro* and *in vivo* and demonstrate their potentials in regulating retinal development. Together, our study presents a rich and

complex landscape of translational regulation and provides novel insights into their roles during retinogenesis.

Keywords: Retinal development, ribosome profiling, translational regulation, microprotein

Introduction

The vertebrate retina is a specialized part of the central nervous system with diverse cell types, high-level organization, and an evolutionarily conserved structure [1]. It can serve as an ideal model to study neural development, such as deciphering the developmental gene regulatory patterns and understanding mechanisms of morphogenesis formation and specificity [2-4]. To date, genome-wide molecular characterization of retinogenesis has been understood using transcriptomic [5,6], epigenomic [7,8], and proteomic [9,10] approaches and has identified many molecules that play important roles in regulating the development of retina.

The central dogma of molecular biology states two major steps during the detailed residue-by-residue transfer process of genetic information: transcription and translation, by which information encoded in DNA flows into RNA via transcriptional regulation and ultimately to proteins via translational regulation [11]. Similar to transcription, translation involves a series of highly temporally orchestrated events directed by *cis*-elements and *trans*-factors [12]. Increasing evidence emphasizes the importance of translation of gene expression [13-15]. Dynamic, tight, and coordinated translational regulation can conduce to the growth of multicellular organisms, particularly during a rapid morphological transition, such as development of red blood cell [16], embryonic stem cell differentiation [17], cortical neurogenesis [18], myogenic differentiation [19], and spermioteleosis [20]. In addition, it is suggested that translational regulation might influence plasticity of visual pathway development and function [21]. Nevertheless, a systematic, genome-wide analysis of gene translation for retinogenesis is lacking at present.

Recently, accumulating evidence has suggested that a fraction of putative small open reading frames (sORFs) within long non-coding RNAs (lncRNAs) are translated to encode

functional micropeptides [22]. For example, translational product of a lncRNA, *Dworf*, is of critical importance in regulating contraction-relaxation cycles in muscle [23]. More recently, some circular RNAs (circRNAs) have been shown to encode bioactive micropeptides, with specific cellular and physiological functions [24,25]. These unexpected findings have further emphasized the complex translational regulation in modulating gene expression. However, it is entirely unclear whether the translation of non-coding RNAs and small ORF-mediated translational control exist during neural development.

Herein, we conducted the first survey of the translational landscape of neural development in mouse. We applied ribosome profiling (Ribo-seq), mRNA sequencing (mRNA-seq), and circRNA sequencing (circRNA-seq) to the developing mouse retina from the embryonic to the adult stages. We found that translation was dynamically uncoupled with transcription, particularly with larger expression divergence before eye-opening. We revealed diverse regulatory changes fulfilling the requirements of gene expression outputs at different developmental stages. We further detected dynamic changes in translational efficiency (TE) and discovered thousands of upstream ORFs (uORFs), fine-tuning gene translation. Surprisingly, we identified hundreds of small ORFs from lncRNAs and circRNAs, which were highly developmental stage specific. We validated their translation *in vivo* and *in vitro* and annotated their functions in retinal development. Overall, our study reveals a widespread and complex landscape of translational regulation in mouse retinal development and provides new insights into the regulation of gene expression, particularly sORF-mediated translation, into retinogenesis and neural development.

Results

Ribosome profiling, mRNA and circular RNA sequencing

To understand the translational regulation of gene expression during mammalian retinogenesis, we performed Ribo-seq and mRNA-seq to generate translatome and transcriptome profiles of mouse retina at six developmental stages, including E13 (embryonic day), E15, P0 (postnatal day), P6, P13, and P21, temporally spanning two major developmental events in retina, birth and eye opening (day11-12) (**Fig. 1A**). We also performed circRNA-seq specifically for

transcriptomic profiling of circRNAs at E15, P0, P6, P13, P21, and M9 (month). All the sequencing experiments were done with two biological replicates. In total, the Ribo-seq, mRNA-seq and circRNA-seq yielded 2.07, 1.08, and 2.33 billion raw reads, with an average of around 81.28, 30.98, and 87.38 million reads per library, respectively (**Table S1**).

Ribosome-protected fragments (RPFs) generated from the Ribo-seq had a typical length range of 25- to 35-nucleotide (nt), tightly distributed around a peak of 29- to 30-nt (**Fig. 1B** and **S1D**), a preference mapped to annotated coding sequences (CDS) and 5' untranslated region (**Fig. S1A**), a strong bias toward the translated frame (**Fig. 1C** and **S1B**), and a characteristic three-nucleotide (3-nt) periodic subcodon pattern (**Fig. 1D** and **S1C**). As expected, these characteristics were absent in the mRNA-seq datasets [26]. There was a high degree of agreement between biological replicates, with an average Pearson's correlation coefficient of 0.981, 0.993 and 0.990 for Ribo-seq, mRNA-seq and circRNA-seq, respectively. Principal component analysis (PCA) showed that the samples had a clear segregation of development stages (**Fig. 1E-G**) and the samples from the same stage were more similar to each other than the ones from the different stages (**Fig. S1E-G**). In addition, expression of some known marker genes in retinal development on our datasets, such as *Otx2*, *Pax6*, and *Neurod4*, well agreed with previous studies (**Fig. S2)** [27-29]. Collectively, these results demonstrated that our experimental data were of high quality.

Figure 1. Overview of gene expression of the developing mouse retina.

Coordination of translational and transcriptional regulation during retinal development

Based on the Ribo-seq and mRNA-seq datasets, we identified an average of 11,150 well-translated protein-coding genes and 12,458 well-transcribed protein-coding genes per stage (**Fig. 2A**, see **Methods**). Most of the well-translated and well-transcribed protein-coding genes were shared across all the stages (**Fig. 2B** and **2C**). The shared genes underwent dynamic expression changes during retinal development (**Fig. S3A** and **S3B**), with 91.4%

(9,294 of 10,172 at transcriptional level) and 76.3% (6,961 of 9,119 at translational level) genes showing significant changes in temporal differential expression (false discovery rate (FDR) < 0.05), including many well-known transcription factors directing retinogenesis such as *Rax*, *Crx*, *Otx2*, *Vsx2*, and *Neurod1* (**Fig. S3C**). Correlation analysis showed that transcriptome and translatome were more different at the early stage but became more similar after eye-opening, suggesting uncoupled changes between transcriptome and translatome during retinal development (**Fig. 2D**).

To capture the overall transcriptional and translational changes in gene expression, we integrated mRNA-seq and Ribo-seq datasets to perform differential expression analysis between adjacent stages (see **Methods**; **Fig. 2E**, FDR < 0.05). We observed that the number of differentially expressed genes (DEGs) gradually increased with development until a peak between P6 and P13, up to a total of 5,753, then followed by a dramatic decline. This pattern was consistent with the progression of retinal differentiation and maturation (**Fig. S3D**). To capture the specific translational changes independent of changes in the transcription, we defined differential translational efficiency genes (DTEGs) (see **Methods**). Based on the different patterns of DEGs and DTEGs, we classified DEGs into four distinct regulatory classes, namely, forwarded, exclusive, buffered, and intensified (see **Methods**; **Fig. 2F** and **Table S3**). The forwarded genes have RPF changes that are explained by the mRNA changes. The exclusive genes have changes in TE without mRNA changes. The buffered genes have changes in TE that offsets the mRNA changes and the intensified genes also have changes in TE that amplifies the mRNA changes.

For these DEGs, we observed that differences in transcription were not always forwarded to the translational level and on average, more than 38% of differentially transcribed genes were translationally buffered or intensified (**Fig. 2G**). Of these, translational buffering was more prominent than translational intensification, which further emphasized the existence of extensive translational regulation that shaped gene expression changes during retinal development. Moreover, translational regulation could also influence gene expression independently, with an average of 517 differential genes found in each comparison of adjacent stages whose changes occurred exclusively at the translational level without underlying mRNA

changes. For instance, between E15 and P6, ribosomal occupancy significantly changed for many well-studied genes related to the functionality of neurons, such as *Dhx36* that helps specific microRNA localize to neuronal dendrite [30], *Eea1* that restores homeostatic synaptic plasticity [31], and *Tk2* that ensures neuronal function *in vivo* [32] (**Fig. 2H**).

Notably, genes in the same regulatory class exhibited developmental stage-specific enrichment of gene ontology (GO) (**Fig. 2I**) while genes in the different regulatory classes exhibited class-specific enrichment of GO (**Fig. 2J**). For instance, "neuron projection organization" which is critical for the establishment of retinal visual function was particularly enriched for the exclusive genes during P6 to P13 and intensified genes during P13 to P21. Overall, these results revealed that both translational and transcriptional regulation had important but different roles in the development of retina.

Figure 2. Transcriptional and translational characterization.

Functional importance of translational regulation

We next attempted to examine the relative contribution of translational and transcriptional regulation for the biological functions important to retinogenesis. We first performed GO enrichment analysis for the above-detected differentially transcribed and translated genes (FDR < 0.01; **Table S4**; see **Methods**) and identified specific functions enriched throughout the development of the mouse retina (**Fig. 3A**). Particularly, many of these functions were associated with synapse formation and synaptic transmission at neuromuscular junctions, such as "visual system development", "postsynapse assembly", "postsynapse organization", and "signal release from synapse". We also observed that some biological functions showed stage-specific enrichment. For instance, "positive regulation of chromosome separation" and "microtubule-based protein transport" were exclusively enriched between P0 and P6 as well as between P6 and P13, respectively.

We next calculated the relative percentages of regulatory classes (that is, forwarded, exclusive, intensified, and buffered) of differential genes for each functional term and performed

a PCA to define their contributions to the overall expression change of each term. The manifestations showed these biologically important functions were attributed to different degrees of translational and transcriptional regulation (**Fig. 3B**). For instance, the "visual system development" was under combinatorial regulation of transcriptional and translational control mechanisms. At the early stage, this function was subjected to the major transcriptional forwarded regulation, but as development progressed, transcriptional regulation gradually weakened and translational regulation, particularly translationally exclusive and intensified regulation, gradually strengthened (**Fig. 3C**). Signal transmission-related function "axonal transport", enriched frequently after birth, was mainly subjected to translational regulation. Specifically, about 62.7% of total contributions in "axonal transport" could be attributed to translational regulation (exclusive + intensified + buffered), of which 72.9% came from exclusive and intensified regulation. Besides these, the stage-specific enriched functions that played important roles in the homeostatic maintenance of mature synapses, such as "microtubule-based transport" and "presynaptic modulation of chemical synaptic transmission", were mainly under translational intensified and exclusive regulation, respectively. Our results showed a rich and complex regulation of gene expression during retinal development.

Figure 3. Dissecting Translational regulation in the developing mouse retina.

Dynamics of translational efficiency and contribution of regulatory uORFs

Due to long half-lives (>2 h) for the majority of eukaryotic mRNAs, regulation of their encoded proteins is achieved by controlling mRNA TEs and protein degradation rates [26]. We detected a total of 5,945 differential translational efficiency (DTE) genes between adjacent stages (see **Methods**). Unsupervised clustering by *k*-means revealed the temporal dynamics of TE that were categorized into seven clusters (**Fig. 4A**). Each cluster was composed of distinct genes with specific biological functions (**Table S5**). Particularly, the 1049 DTE genes in the cluster C were significantly enriched in "synapse organization", "postsynapse organization", and "neuron

projection organization". TEs of these genes in the cluster C were specifically enhanced around E15 (embryonic wave) and P13 (postnatal wave), which would help promote neuronal differentiation, thereby facilitating production of functionally active neurons. The 925 DTE genes in the cluster E were significantly enriched in "mRNA/tRNA 5'-end processing", "midbody abscission", and "microtubule-based transport". Notably, these genes showed peak TE at P6 and might play a critical role in maintaining retina homeostasis and later neurogenic divisions, given typical characteristics of extensive alternative splicing [33] and active cell division events [34] during this period of retinal development. In addition, the 816 DTE genes in the cluster B showed peak TE at the embryonic period and were mainly enriched in "mitotic nuclear division", "negative regulation of chromatin organization", and "positive regulation of chromatin organization", fulfilling the requirements for early neurogenic divisions.

Previous studies have shown that regulatory elements located on mRNA transcripts might affect gene translation [35], particularly upstream open reading frames (uORFs) in the 5'-UTRs [36]. Therefore, we detected a total of 2,971 actively translated uORFs in 2,123 protein-coding genes (**Fig. 4B** and **Table S6**). Comparing with mass spectrometry (MS)-based proteomics, we provided direct *in vivo* protein evidence for translation of 181 uORFs (**Table S6**, see **Methods**). In addition, we randomly selected four uORFs and successfully validated their translated products by *in vitro* translation (IVT) assays (**Fig. 4C**).

We observed significant differences of TEs of genes with and without uORFs, which showed uORF-mediated translational repression of main ORFs at the early stage (**Fig. 4D**). However, after birth, the uORF-mediated translational repression of main ORFs did not always occurred, which could be achieved by ribosome bypass or translation re-initiation [37]. The majority (57.9%, 1,719/2,971) of uORFs were found exclusively in a single stage, implying that uORF-mediated translational regulation occurred in a stage-specific manner (**Fig. 4E**). The GO enrichment analysis revealed that uORF-containing genes participated in many important biological processes required for retinal development, such as "axon guidance", "dopaminergic synapse", and "signaling pathways regulating pluripotency of stem cells" (**Fig. 4E** and **Table S7**). In addition to thousands of translated uORFs, only 266 downstream ORFs (dORFs) were detected in the 3'-UTRs of 217 protein-coding genes, with an average of 64 dORFs (57 genes)

per stage. However, by comparing TEs of genes with and without dORFs, we did not observe that translation of dORFs significantly enhances translation of their corresponding canonical ORFs, as reported in a previous study [38] (**Fig. S4A**).

Figure 4. Analysis of translational efficiencies across retinal development.

Translation of long noncoding RNA genes

Microproteins encoded by presumed long noncoding RNAs (lncRNAs) have frequently been overlooked, and their prevalence and potential function in retinal development, even in neural development, remain unknown. To discover translated lncRNAs in the developing mouse retina, we searched for actively translated ORFs in lncRNAs (lncORFs) (see **Methods**). In total, we identified 603 unique lncORFs in 290 lncRNAs, with a median length of 48 amino acids (aa) per lncORF (**Fig. 5A** and **Table S8**). Majority of these lncORFs (57.5%, 347/603) could be found in another translatome dataset of mouse retina (GSE94982), including previously well-characterized microproteins *Pax6os1* [39], *B230354K17Rik* [40], and *Crnde* [24]. Comparing with MS-based proteomics, we provided direct *in vivo* protein evidence for translation of 75 out of the 290 lncRNAs (**Table S8**, see **Methods**).

To further experimentally validate the translation products, we first performed *in vitro* translation (IVT) assays for 10 randomly chosen translated lncRNAs. Considering the relatively low molecular weight of potential microproteins, we specifically fused a 11-amino-acid HiBiT epitope tag to the C-terminal of each lncORF that could produce bright and quantitative luminescence through high affinity complementation with LgBiT [41]. The quantifiable results demonstrated that all of them successfully produced microproteins in IVT assays (**Fig. 5B**). Furthermore, we separately transfected the expression vectors of these lncORFs into N2a (Neuro-2a), Hela, and ARPE19 cells and detected microprotein products of all of them in at least two cell types (**Fig. 5C**). Subsequent start codon mutation of these lncORFs prevented their translation or caused truncated translation, as evidenced by significant decreases in the

luminescence intensity (**Fig. 5D** and **S5A**). Loss of signal in the predicted size range of microprotein products was additionally confirmed by western blot analysis (**Fig. 5E** and **S5B**).

Translation of lncRNAs showed a strong stage-specificity, with an average of 112 lncRNAs per stage detected to undergo active translation but only 19 shared across all developmental stages (**Fig. S5C**). Of the 290 translated lncRNAs, 198 exhibited significant temporal expression changes (**Fig. S5D**), suggesting that most of translation events of lncRNAs independently contributed to changes in developmental programs (**Fig. S5E**). We also sought to understand the tissue-specificity of these 290 translated lncRNAs. We conducted Ribo-seq and RNA-seq for six tissues in mouse (GSE94982) and found among the 290 translated lncRNAs, 172 (59.3%), 102 (35.2%), 133 (45.9%), 79 (27.2%), and 128 (44.1%) that also had translation evidence in the brain, heart, kidney, liver, and lung, respectively, with 18 translated in all six tissue types. Importantly, the majority of these lncRNAs (86.1%-91.5%) shared at least one identical lncORF with those detected in the retina. These results suggested that some of the translated lncRNAs detected in the developing retina had potential multi-tissue activities.

Functional annotation of these microproteins translated from lncRNAs using InterProScan 5 revealed 68 (~11.3%) of 603 lncORFs with identifiable features found in known proteins (**Table S9**), such as conserved domains, protein families and functional sites, some of which even could be assigned to defined molecular functions. For instance, microprotein encoded by *Ptpmt1* participated in “protein tyrosine/serine/threonine phosphatase activity” and microprotein encoded by *RP23-95L9.6* participated in “nucleic acid binding”. Our analysis suggested that these lncORFs had functional relevance which could be tested in the future study.

Figure 5. Translation of lncRNAs.

Translation of circRNAs

In addition to lncRNA translation, an increasing number of studies have indicated that circRNAs can also be translated into detectable peptides with physiological functions [42]. We used circRNA-seq and Ribo-seq data to discover translation potentials of circRNAs. Using a

stringent identification strategy, 262 of 28,910 transcribed circRNAs with at least two supporting back-spliced footprints were defined as ribosome-associated circRNAs (ribo-circRNAs) (**Fig. 6A** and **S6A**, see **Methods**). Notably, the majority (83.2%) of them were found in the stages after eye opening, suggesting their potential functions in the later stage of retinal development (**Fig. 6B**).

Ribo-circRNAs had distinguishing properties from untranslated circRNAs (ut-circRNAs). Ribo-circRNAs had significantly shorter exonic length and fewer exons than the ut-circRNAs, with at least 29% and 40% decreases in the median length and number of exons, respectively (Wilcoxon rank-sum test, $P < 2.2e-16$; **Fig. 6C**), in agreement with previous observations in human cell lines [43]. Ribo-circRNAs had significantly longer flanking introns (Wilcoxon rank-sum test, $P = 2.88e-13$) that harbored more repetitive elements than ut-circRNAs (Wilcoxon rank-sum test, $P = 1.04e-13$, **Fig. 6D-E**). In addition, evolutionary conservation of microproteins translated from ribo-circRNAs was significantly higher than that of ut-circRNAs (Wilcoxon rank-sum test, $P < 1.94e-03$, **Fig. 6F**), suggesting that they were likely to have functional roles, given that strong conservation in sequence is a general indicator of important biological function [44].

Next, we randomly selected 20 candidates from 262 ribo-circRNAs and performed Sanger sequencing of RT-PCR products using divergent primers to confirm the back-splice junction sites of these circRNAs (**Fig. 6G** and **S6C**). After RNase R treatment, all these circRNAs were resistant to digestion with RNase R exonuclease, validating the existence of these ribo-circRNAs (**Fig. S6B**). Comparing with MS-based proteomics, we further provided direct *in vivo* protein evidence for 13 out of 262 ribo-circRNAs (**Table S10**). Of the 13 circRNAs producing detectable microproteins, six were found to be differentially transcribed, including *circKmt2a*, *circArhgap10*, *circMacf1*, *circAp3b2* (up-regulated in P21 compared to P6), *circWdr7* (up-regulated in P0 compared to E15), and *circZfp532* (down-regulated in P21 compared to P6), emphasizing that their translation might be important for retinal development.

Figure 6. Translation of circRNAs.

Discussion

Our integrative analysis of transcriptome and translatome during retinal development revealed specific changes in translation and regulatory roles of translational regulation in retinogenesis. We found that retinogenesis is accompanied by dynamic, rapid and coordinated changes in gene translation and translational regulation. Thousands of important regulatory protein-coding genes were subjected to significant changes in the translational level in a stage-specific manner, potentially redefining functional architecture and diversity of the retinal cells.

Specific translational regulation could be achieved by controlling TEs, thereby determining quantitative differences in protein abundance during retinogenesis. TE dynamics during developmental transition could further be triggered by some regulatory elements, such as the most commonly used uORFs. Within 5'-UTRs, uORF-mediated translational control was a vitally regulatory mechanism for gene expression in mammals [45]. Our findings revealed that uORFs provided functionally important repression for many key genes associated with retinal development, such as an uORF existing within *Otx2* gene only at P0 and two uORFs existing within *Nrl* gene at E13, P0, P6 and P13, displaying stage-specific influence on retinogenesis. Moreover, uORFs exerted a substantial repressive effect on translation of primary CDS in a dose-dependent manner, demonstrating the complexities of uORF-mediated translational regulation [46].

Surprisingly, we discovered pervasive translation of non-canonical ORFs, including those from lncRNAs and circRNAs. We experimentally confirmed their translation products and showed that these micropeptides had regulatory potential and biological relevance. For instance, we found that microprotein encoded by *RP23-15O6.3* contained C2H2-type zinc fingers, and that microprotein encoded by *AC166710.3* shared a common structure with ribosomal L18e/L15P superfamily. Functional enrichment analysis revealed potential roles of these translated lncRNAs in retinal development. One representative lncRNA was *Miat* (also known as *Rncr2* or *Gomafu*) (see **Methods** and **Fig.S5D**) contributed to mitosis of retina progenitor cells, illustrating its essential role for retinogenesis [49,50]. Notably, besides the direct involvement of retinal development, some of the encoded microproteins may serve as

regulatory elements to play important roles in modulating neighboring gene activity. This is partially supported by the observation that neighboring genes of many translated lncRNAs such as *Six3os1* (IVT assays and *in vivo* translation), *Vax2os, Otx2os1, Pax6os1*, and *Zeb2os1* (MS detection) were important determinations of retinal cell fate [51]. In addition, 13 microproteins encoded by circRNAs were detected during retinal development. These microproteins might be partially associated with the functions of their host proteins due to the existence of overlapping sequences between them, likely competing with [52] or protecting [53] their host proteins. Notably, in our study, microproteins were only evident for a relatively small subset of ribosome-associated lncRNAs and circRNAs. One reason was that some microproteins were not abundant enough to be detected by the current technologies [54]. And another reason is that some microproteins were unstable and rapidly degraded [55]. To obtain a comprehensive map of hidden microproteins during retinal development, innovative technologies are needed.

To sum, our study provides a comprehensive view and important biological insights how translational regulation instructs retinal development and our dataset serves as a valuable resource for future studies of the translational machinery during retinogenesis.

Materials and methods

Animals and tissue collection

Wild-type mice of C57BL/6J genetic background were purchased from Guangdong Medical Experimental Animal Center. Mice from 7 different developmental stages (including E13, E15, P0, P6, P13, P21 as well as M9) were euthanized and the eyes were enucleated immediately after sacrifice. Retinal tissues from enucleated eyes of each mouse were harvested and then snap-frozen with liquid nitrogen for next-generation sequencing. All experimental procedures were approved by the Animal Ethics Committee of Zhongshan Ophthalmic Center, Sun Yat-sen University (Guangzhou, China; License No: SYXK (YUE) 2018-0189).

Library preparation and sequencing

For mRNA sequencing and ribosome profiling, frozen samples were lysed using 1 ml of mammalian lysis buffer (200 µl of 5x Mammalian Polysome Buffer, 100 µl of 10% Triton X-100, 10 µl of DTT (100 mM), 10 µl of DNase I (1 U/µl), 2 µl of cycloheximide (50 mg/ml), 10 µl of 10% NP-40, and 668 µl of nuclease-free water). After incubation for 20 minutes on ice, the lysates were cleared by centrifugation at 10,000xg and 4°C for 3 minutes. For each tissue and replicate sample, the lysate was divided into 300-µl and 100-µl aliquots. For the 300-µl aliquots of clarified lysates, 5 units of ARTseq Nuclease were added to each A260 lysate, and the mixtures were incubated for 45 minutes at room temperature. Nuclease digestion was stopped by additional 15 µl of SUPERase·In RNase Inhibitor (Ambion). Subsequently, the lysates were applied to Sephacryl S400 HR spin columns (GE Healthcare Life Sciences), and ribosome-protected fragments were purified using the Zymo RNA Clean & Concentrator-25 kit (Zymo Research). Ribosomal RNA was depleted using the Ribo-Zero magnetic kit (Epicentre). Sequencing libraries of ribosome protected fragments were generated using the ARTseq™ Ribosome Profiling Kit (Epicentre, RPHMR12126) according to the manufacturer's instructions. From the 100-µl aliquots of clarified lysates, poly(A)+ RNAs were extracted and purified, and sequencing libraries of poly(A)+ RNAs were then generated using the VAHTSTM mRNA-seq v2 Library Prep Kit from Illumina (Vazyme Biotech, NR601-01) according to the manufacturer's instructions. The resulting 26 barcoded libraries were pooled and sequenced using an Illumina HiSeq 2500 instrument in the single-end mode.

For circRNA sequencing, total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA, USA), the RNA integrity was assessed by Agilent 2100 with RIN number >7.0. Approximately 5 µg of total RNA was used for library preparation, the Ribo-Zero™ rRNA Removal Kit (Illumina, San Diego, USA) was used to deplete ribosomal RNA according to the manufacturer's instructions, the left RNAs were treated with Rnase R (Epicentre Inc, Madison, WI, USA) to remove linear RNAs and to enrich circRNAs. The circRNA sequencing libraries were then generated using TruSeq Stranded Total RNA HT Sample Prep Kit (Illumina RS-122-2203) according to the manufacturer's instructions with minor modifications. Briefly, cDNAs were reverse transcribed using ProtoScript II Reverse Transcriptase. The ligated

products were amplified with PCR by the following conditions: initial denaturation at 95˚C for 3 min; 8 cycles of denaturation at 98˚C for 15 sec, annealing at 60˚C for 15 sec, and extension at 72˚C for 30 sec; and then final extension at 72℃ for 5min. At last, the resulting 10 barcoded libraries were sequenced using an Illumina Hiseq X Ten instrument in the paired-end mode.

mRNA-seq and Ribo-seq alignment

The raw ribosome profiling and mRNA-seq sequencing reads were demultiplexed using CASAVA (version 1.8.2), followed by adapter trimming with Cutadapt (v1.9.1, -e 0.1 -O 6 -m 20) [56] and removal of poor-quality reads with Sickle (v1.33, -q 20 -l 20 -x -n) [57]. The reads aligned to mouse tRNA and rRNA sequences using Bowtie (v1.0.1, -l 20) [58] were further removed. All remaining reads were mapped to the mouse reference genome (GENCODE, Release M18: GRCm38.p6) using STAR (v2.5.2) with default parameters. Only uniquely mapped reads selected by samtools (v1.6, Phred score >= 20) were used for subsequent analysis.

Gene quantification and expressed genes definition

Raw counts for different genomic features were obtained using featureCounts from the subread package in Bioconductor (v1.5.0, -t CDS for ribosome profiling data and –t exon for mRNA-seq data) [59]. The raw counts from all 26 samples were then combined and normalized together by a pool-based size factor yielded from DESeq2 R package [60] to minimize the batch effects among samples (**Table S2**). After that, the expression level of each gene was estimated as reads per kilobase of transcript per million reads mapped (RPKM) using an in-house R script. At the transcriptome layer, only those genes with RPKM > 1 across two replicated samples of each developmental stage were kept, defined as well-transcribed genes. At the translatome layer, only those genes that are well-transcribed and undergo active translation (see below) were defined as well-translated genes.

Actively translated ORF detection

Actively translated ORF detection in ribosome profiling data was performed using RiboTISH (v0.2.4) with '--framebest' strategy to select the best candidate ORF [61]. To increase statistical

power of ORF finder, we merged the aligned BAM files from two replicated samples of each developmental stage using 'samtools merge' (v1.6). uORFs were defined as ORFs originating from the 5'-UTRs of annotated protein-coding genes (that is, with TisType: 5'-UTR), dORFs were defined as ORFs originating from the 3'-UTRs of annotated protein-coding genes (that is, with TisType: 3'-UTR), and lncORFs were defined as ORFs originating from annotated long noncoding RNA genes.

Genes with dynamical changes detection

Genes with dynamic expression patterns over time were identified by the maSigPro R package, which is specifically designed for the analysis of time-course gene expression data [62]. Genes with significant temporal changes in retinal development were selected at a FDR < 0.05 and R-squared threshold equal to 0.6.

Translational efficiency Estimation

The TE for each protein-coding gene was estimated as the ratio of the normalized TPM values of Ribo-seq to RNA-seq reads in annotated CDS regions [63]. Given a high degree of TE correlation between two replicated samples of each developmental stage (mean Pearson's correlation coefficient = 0.911), TE values of the two replicates for each protein-coding gene were then averaged to form a final TE for subsequent comparative analysis.

Differential expression analysis, gene classification, and gene ontology (GO) enrichment analysis

Differentially transcribed and translated genes were detected using deltaTE [64], which incorporates Ribo-seq and mRNA-seq data. Genes with change in mRNA and RPF levels at the same rate were defined as differentially transcribed genes (DTG), and genes with change in RPF level independent of change in mRNA level, which lead to a change in TE, were defined as differential translational efficiency gene (DTEG). DTGs and DTEGs between adjacent stages could further be categorized into four classes: buffered, intensified, forwarded and exclusive. Specifically, translationally buffered genes have a significant change in TE that offsets the change in RNA; translationally intensified genes have a significant change in TE that bases on

the effect of transcription; translationally forwarded genes are DTGs that have a significant change in mRNA and RPF at the same rate and with no significant change in TE. Conversely, translationally exclusive genes are DTEGs that have a significant change in RPF but with no change in mRNA leading to a significant change in TE.

GO enrichment analysis was used to reveal biological functions of differentially expressed genes (DEGs), which was achieved by the ClusterProfiler R package [65]. Only those GO terms with false discovery rate (FDR) < 0.01 were regarded as statistically significant.

Principal component analysis

To define the main contributing layer of gene expression regulation (buffered, exclusive, forwarded, and intensified) for each coregulatory biological function, principal component analysis (PCA) was performed as described in a previous study [24]. For each GO term, the relative fractions of four defined classes of differential genes were used as an input for the PCA. The prcomp and fviz_pca_biplot functions from the factoextra R package were used for the PCA and visualizing the output of the PCA, respectively [66]. Thus, the placement of each GO term in the PCA plot was based on the directionality of four layers of gene expression regulation.

CircRNA identification and quantification

Raw circRNA-seq reads were preprocessed with Perl scripts, including the removal of adaptor-polluted reads, low-quality reads (Phred score >= 20) and reads with number of N bases accounting for more than 5%. The clean reads were then mapped to the mouse reference genome (GENCODE, GRCm38.p6) using BWA-MEM (-T 19, v0.7.17) [67]. Next, two different detection tools were used to identify transcribed circRNAs, namely, CIRI2 (v2.0.6) [68] and CIRCexplorer2 (v2.3.6) [69]. To further reduce false positives of circRNA identification, only those circRNAs meeting all of the following three criteria were kept, including (1) having at least 2 unique backsplice junction (BSJ) reads, (2) being simultaneously identified in both tools, and (3) being simultaneously detected in two replicated samples of each developmental stage. Finally, the BSJ reads of CIRI2 were used to quantify the transcriptional level of circRNA using CPM (counts per million mapped reads) [70].

Identification of Ribosome associated-circRNAs

To determine ribosome-associated circRNAs, we first extracted the 40-base pair (bp) sequences on either side of the backsplice junction site of each transcribed circRNA, and then the sequence was ligated in tandem to generate a pseudo circRNA reference. Next, all Ribo-seq reads that failed to map to the linear reference genome were realigned to the pseudo circRNA reference sequences using Tophat2 (v 2.1.1) with default parameters except N, which was set to 0 (the default is 2) [71]. Finally, ribosome associated-circRNAs (ribo-circRNAs) were defined as having (1) at least one unique backsplice junction-spanning Ribo-seq reads and (2) a minimum read-junction overlap of three nucleotides (nt) on either side of the backsplice junction site.

cORF prediction

To predict putative circRNA-encoded ORF (cORF), the cORF_prediction_pipeline was used with some modifications. Briefly, all possible cORFs with an AUG-start codon followed by an in-frame stop codon in the exonic sequence were identified and then filtered based on the requirements of a minimum length of 20 amino acids (aa) and of spanning the backsplice junction site. Notably, only the longest cORF was retained for each of the three frames per ribo-circRNA. If a cORF does not contain stop codon, it was defined as an INF (infinite)-cORF, representing that the circRNA could be translated via a rolling circle amplification mechanism.

Two-dimensional liquid chromatography-tandem mass spectrometry

Retina tissues at E13, E15, P0, P6, P13 and P21 were mixed, fully ground in liquid nitrogen and lysed with lysis buffer L3 (Fitgene Biotech, #FP1801), 0.2% SDS and 1x PMSF (Sangon Biotech, #P0754). Lysates were sonicated on ice and cleared by centrifugation at 12000 rpm for 10 min. The protein extracts were purified by overnight acetone precipitation and re-solubilized using L3, concentration of the samples was determined by Bradford assay. Proteins were reduced using 50 mM TCEP (1 hour at 60˚C), alkylated using 55 mM MMTS (45 min at RT in the dark), then the protein sample was loaded on a 10 kDa ultrafiltration tube, washed twice with 8M Urea and three times with 0.25M TEAB. For protein digestion, 50 μl of

0.5M TEAB and trypsin (Promega, enzyme: protein ratio of 1:50) were added to the membrane, reaction was incubated overnight at 37˚C. The resulting peptides were first fractionated on a Gemini-NX 3u C18 110Å 150*2.00 mm columns (Phenomenex, #00F-4453-B0) using high-pH reversed-phase chromatography (Dionex Ultimate 3000 RSLCnano) with increasing concentration of acetonitrile, 20 fractions were collected according to the 214 nm absorbance and running time. After vacuum drying, 3 µg of fractionated peptide samples were separated on an Acclaim PepMap RSLC C18 2µm 100Å 75 µm i.d. x 150mm column (Dionex, #160321) and analyzed using Thermo Scientific™ Q Exactive™. Full MS spectra from m/z 375-1800 were acquired at a resolution of 70,000 with an automatic gain control (AGC) target value of 3e6 and maximum injection time (IT) of 40 ms. MS/MS spectra were obtained at a 17,500 resolution with an AGC target of 1e5 and maximum injection time (IT) of 60 ms, TopN was set to 20 and NCE/stepped NCE was set to 27. In addition, two unfractionated peptide samples were also analyzed by LC-MS/MS using the same parameters as fractionated peptide samples.

Analysis of mass spectrometry-based proteomic data

Three publicly available proteomic datasets obtained from the PRIDE database (Accession number: PXD003441, PXD002247, and PXD009909) and an in-house MS dataset were used to detect protein products from translatable lncRNAs and circRNAs. The raw data files were analyzed using MaxQuant software (v1.6.15.0) [72] against a custom-made database, which combined all mouse sequences from UniProt/Swiss-Prot (MOUSE.2020-08) with sequences derived from translatable lncRNAs and circRNAs, based on the target decoy strategy (Reverse) with the standard search parameters with the following exceptions: (1) the peptide-level FDR was set to 5%, and the protein-level FDR was excluded; (2) the minimal peptide length was set to 7 amino acids; and (3) a maximum of two missed cleavages was allowed. Each search included carbamidomethylation of cysteine as fixed modification methionine oxidation, N-terminal acetylation as variable modifications, but for PXD002247 and our own MS data, a variable modification of deamidation of asparagine and glutamine was also included. Finally, a total of 74 lncRNA-derived and 13 circRNA-derived peptides were evidenced by at least one unique peptide.

Functional annotation of lncORF-encoded peptides

Conserved domain and protein homology detection. Each of the putative peptides encoded by lncORFs (SEPs) was annotated against the InterPro database using InterProScan (v5.44) with the default parameters [73]. In total, 212 of 603 SEPs were mapped to known homologous records in the InterPro database, of which 68 were annotated with specific functional domains involving in vital important pathways (**Table S9**).

Guilt–By–Association Approach. A total of 198 translatable lncRNAs with dynamical changes in the retinal development were grouped into three clusters by using k-means clustering algorithm. To infer biological functions of each lncRNA cluster, the guilt–by–association approach was used, where (1) Pearson's correlation coefficient between each lncRNA-mRNA pair was computed in all samples in our dataset; (2) candidate lncRNA-mRNA pairs were selected as those with correlation coefficients > 0.70 and significance level of 0.05 for Pearson's correlation (FDR<0.05); and (3) protein-coding gene coexpressed with any one lncRNA of each cluster were merged into a union set of genes for GO enrichment analysis.

Translation validation of uORFs and lncORFs

Plasmid construction. We constructed a series of expression vectors for the detection of non-canoncial ORF translation, including 4 uORFs (*u-Rnf10, u-Rnft1, u-Usp8* and *u-Zkscan17*) and 10 lncORFs (*Brip1os, Cct6a, Gas5, Malat1, Miat, Peg13, RP23-41oL16.1, RP24-112I4.1, Six3os1*, and *Zfas1*), together with two known translatable lncRNAs (*Mrln* and *Dworf*) and one protein-coding gene (*DHFR*) as positive controls. The uORFs and lncORFs (including the predicted 5'- and 3'-UTR) were cloned from mouse retinal cDNAs; *Mrln* were cloned from mouse muscle cDNAs; *Dworf* were cloned from mouse heart cDNAs; and *DHFR* were cloned from Hela cells cDNAs. The primers used for cloning were listed in **Table S11**. After obtaining the sequence of these candidates and controls, a HiBiT tag was inserted upstream the stop codon of each predicted ORF to enable luminescence-based detection of translation products. The HiBiT tags were added by inverse PCR using KOD -Plus- Mutagenesis Kit (TOYOBO, #SMK-101), and the primers used for inverse PCR were listed in **Table S11**. In order to confirm that peptides were indeed translated from the corresponding ORFs, mutations were introduced

in the start codons of predicted ORFs using ClonExpress II One Step Cloning Kit (Vazyme, #C112-02). Additionally, 10 ORFs (*Cct6a, Gas5, Malat1, Mrln, RP23-41oL16.1, RP24-112I4.1, Six3os1, Zfas1, u-Rnft1 and u-Usp8*) and their mutants were added with 3xFlag tags downstream the HiBiT tag. The primers and oligonucleotides used for mutation generation and 3xFlag insertion were listed in **Table S11**. Sequences of all constructed plasmids were verified by sanger sequencing and the plasmid DNA was extracted using EndoFree Plasmid Midi Kit (CWBIO, #CW2105).

Luminescence-based detection of translation products. The translation of ORFs were detected both in IVT assays and cultured cells. For IVT assays, TnT® Quick Coupled Transcription/Translation System was used for *in vitro* translation of all HiBiT-tagged ORFs. According to the manufacturer's instructions, 1 μg of plasmid DNA was used as template for each reaction and the reaction was incubated for 90 min at 30˚C. The detection of translated products was performed using Nano-Glo® HiBiT Lytic Detection System (Promega, #N3030) according to the manufacturer's instructions with minor modifications. In Brief, 10 μl IVT products were diluted to 50 μl using nuclease free water, added with 50 μl lytic buffer and mixed by pipet, after 10 minutes incubation at room temperature, the luminescence was measured on a Promega GloMax®-Multi Detection System. To further explore the translational potential of the candidates in cellular context, 500ng of plasmid DNA was transfected into N2A, Hela and ARPE19 cells using Lipofectamine 3000. Cells were harvested at 48 hr post-transfection for subsequent analysis. The plate was equilibrated to room temperature, then 300 μl lytic buffer was added to each well and incubated for 10 minutes on orbital plate shaker. The lysates were divided into three tubes and the luminescence of each tube was measured on a Promega GloMax®-Multi Detection System.

Western blots. The translated peptides were further validated by western blot. Transfected N2A Cells were harvested with 100 μl of RIPA buffer (25 mM Tris•HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) added with 1X PIC (Merck, #539131) and Benzonase (NovoProtein, #M046-01B) and incubated 10 min on ice. For detection of peptides from *Brip1os*, *Miat*, *u-Rnf10* and *DHFR* , lysates were added with 5×SDS-PAGE Sample Buffer (GenStar, #E153) and denatured at 95˚C for 5 minutes, and samples were loaded on 4–20%

Mini-PROTEAN® TGX™ Precast Protein Gel and transferred to a 0.2 μm NC membrane. For detection of the remaining peptides, lysates were added with 2x Novex Tricine SDS Sample Buffer (Invitrogen, #LC1676) and denatured at 85˚C for 2 minutes, and samples were loaded on 16.5% GLASS Gel®Tricine gel (WSHT, #TCH2001-16.5T) and transferred to a 0.1μm NC membrane. The blot was carried out using Nano-Glo® HiBiT Blotting System (Promega, #N4210) according to the manufacturer's instructions.

Validation of circRNAs

Divergent PCR validation. Divergent primers were designed to amplify the circRNA backsplice junction sequence and retinae cDNA were used as template for divergent PCR (**Table S11**). Divergent PCR was performed using green Taq mix (Vazyme), and the reaction was carried out for 3min at 95˚C and 30 cycles of 15 sec at 95˚C, 15 sec at 60˚C and 30 sec at 72˚C. The PCR products were then analyzed on 1.5% agarose gels in 1x TAE buffer. All of the PCR products were sanger sequenced with forward and reverse primers to find the backsplice junction sequence.

RNase R treatment assay. RNase R treatment and circRNA quantification was performed according to a published protocol (Panda and Gorospe, 2018) with minor modification. In brief, RNA was treated with 20 μl RNase R digestion reaction (2 μg RNA, 1 μl RNase R (Lucigen, #RNR07250), 2 μl 10x RNase R reaction buffer) and control reaction without RNase R. The reactions were incubated for 30 min at 37˚C and immediately purified using ZYMO RNA Clean & Concentrator (ZYMO RESEARCH, D7011). The purified RNA samples were eluted in 20 μl of nuclease free water, and 12 μl of RNase R treated RNA and control RNA were used for reverse transcription. Quantification of circRNAs were performed using iTaq Universal SYBR Green Supermix (BioRad, #1725124) according to the manufacturer's instructions. The primers used for circRNA quantification were listed in **Table S11** and qPCR reactions were prepared as follows: 0.1 μl cDNA, 10 μl of SYBR Green Supermix, 1.2 μl primer mix (5 μM each) and 6.8 μl nuclease-free water. Reactions were carried on CFX Connect Real-Time PCR Detection System for 2 min at 95˚C and 40 cycles of 5 sec at 95˚C and 20 sec at 60˚C followed by melting

curve analysis. The enrichment of RNA after RNase R treatment was calculated using the delta CT method, mouse gene *Gapdh* and *Rps14* were used as linear control.

Data Availability

All Ribo-Seq, RNA-Seq, and circRNA-seq sequencing data have been deposited in the SRA database at NCBI (https://www.ncbi.nlm.nih.gov/sra/) under accession number PRJNA589677. The mass spectrometry proteomics data have been deposited in the PRIDE database under accession number PXD023439.

Acknowledgments

We thank the support for the Center for Precision Medicine, Sun yat-sen University.

Author Contributions

Z.X. and H.W.W. supervised the project; J.Q.Y. performed RNA-seq and Ribo-seq library construction; C.Y.C. designed and performed experiments; M.Q.X provided help and suggestions for experimental design; K.N.C and H.H.L. analyzed and interpreted the data; K.N.C., C.Y.C., H.W.W., and Z.X. wrote the manuscript; All authors approved the manuscript.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Supplementary Materials

Supplementary materials can be found online.

Funding

This work was supported by the National Natural Science Foundation of China [31871302 to Z.X.] and the Joint Research Fund for Overseas Natural Science of China [31829002 to Z.X.].

References

1. Zaghloul NA, Yan B, Moody SA. Step-wise specification of retinal stem cells during

normal embryogenesis. *Biology of the cell* 2005; **97**(5): 321-37.

2. Bassett EA, Wallace VA. Cell fate determination in the vertebrate retina. *Trends in neurosciences* 2012; **35**(9): 565-73.

3. Holguera I, Desplan C. Neuronal specification in space and time. *Science (New York, NY)* 2018; **362**(6411): 176-80.

4. Abdusselamoglu MD, Eroglu E, Burkard TR, Knoblich JA. The transcription factor odd-paired regulates temporal identity in transit-amplifying neural progenitors via an incoherent feed-forward loop. *eLife* 2019; **8**.

5. Kim JW, Yang HJ, Oel AP, et al. Recruitment of Rod Photoreceptors from Short-Wavelength-Sensitive Cones during the Evolution of Nocturnal Vision in Mammals. *Developmental cell* 2016; **37**(6): 520-32.

6. Lu Y, Shiau F, Yi W, et al. Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. *Developmental cell* 2020; **53**(4): 473-91.e9.

7. Aldiri I, Xu B, Wang L, et al. The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron* 2017; **94**(3): 550-68.e10.

8. Norrie JL, Lupo MS, Xu B, et al. Nucleome Dynamics during Retinal Development. *Neuron* 2019; **104**(3): 512-28.e11.

9. Barnhill AE, Hecker LA, Kohutyuk O, Buss JE, Honavar VG, Greenlee HW. Characterization of the retinal proteome during rod photoreceptor genesis. *BMC research notes* 2010; **3**: 25.

10. Baudet ML, Hassanali Z, Sawicki G, List EO, Kopchick JJ, Harvey S. Growth hormone action in the developing neural retina: a proteomic analysis. *Proteomics* 2008; **8**(2): 389-401.

11. Crick F. Central dogma of molecular biology. *Nature* 1970; **227**(5258): 561-3.

12. Kong J, Lasko P. Translational control in cellular and developmental processes. *Nature reviews Genetics* 2012; **13**(6): 383-94.

13. Schafer S, Adami E, Heinig M, et al. Translational regulation shapes the molecular landscape of complex disease phenotypes. *Nature communications* 2015; **6**: 7200.

14. Tahmasebi S, Amiri M, Sonenberg N. Translational Control in Stem Cells. *Frontiers in*

genetics 2018; **9**: 709.

15. Teixeira FK, Lehmann R. Translational Control during Developmental Transitions. *Cold Spring Harbor perspectives in biology* 2019; **11**(6).

16. Alvarez-Dominguez JR, Zhang X, Hu W. Widespread and dynamic translational control of red blood cell development. *Blood* 2017; **129**(5): 619-29.

17. Atlasi Y, Jafarnejad SM, Gkogkas CG, Vermeulen M, Sonenberg N, Stunnenberg HG. The translational landscape of ground state pluripotency. *Nature communications* 2020; **11**(1): 1617.

18. Blair JD, Hockemeyer D, Doudna JA, Bateup HS, Floor SN. Widespread Translational Remodeling during Human Neuronal Differentiation. *Cell reports* 2017; **21**(7): 2005-16.

19. de Klerk E, Fokkema IF, Thiadens KA, et al. Assessing the translational landscape of myogenic differentiation by ribosome profiling. *Nucleic acids research* 2015; **43**(9): 4408-28.

20. Wang ZY, Leushkin E, Liechti A, et al. Transcriptome and translatome co-evolution in mammals. *Nature* 2020; **588**(7839): 642-7.

21. Starr CR, Pitale PM, Gorbatyuk M. Translational attenuation and retinal degeneration in mice with an active integrated stress response. *Cell death & disease* 2018; **9**(5): 484.

22. Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* 2015; **4**: e08890.

23. Nelson BR, Makarewich CA, Anderson DM, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science (New York, NY)* 2016; **351**(6270): 271-5.

24. van Heesch S, Witte F, Schneider-Lunitz V, et al. The Translational Landscape of the Human Heart. *Cell* 2019; **178**(1): 242-60.e29.

25. Koch L. RNA: Translated circular RNAs. *Nature reviews Genetics* 2017; **18**(5): 272-3.

26. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011; **147**(4): 789-802.

27. Cepko C. Intrinsically different retinal progenitor cells produce specific types of progeny. *Nature reviews Neuroscience* 2014; **15**(9): 615-27.

28. Ren Q, Yang CP, Liu Z, et al. Stem Cell-Intrinsic, Seven-up-Triggered Temporal Factor

Gradients Diversify Intermediate Neural Progenitors. *Current biology : CB* 2017; **27**(9): 1303-13.

29. Xiang M. Intrinsic control of mammalian retinogenesis. *Cellular and molecular life sciences : CMLS* 2013; **70**(14): 2519-32.

30. Bicker S, Khudayberdiev S, Weiß K, Zocher K, Baumeister S, Schratt G. The DEAH-box helicase DHX36 mediates dendritic localization of the neuronal precursor-microRNA-134. *Genes & development* 2013; **27**(9): 991-6.

31. Xu X, Pozzo-Miller L. EEA1 restores homeostatic synaptic plasticity in hippocampal neurons from Rett syndrome mice. *The Journal of physiology* 2017; **595**(16): 5699-712.

32. Bartesaghi S, Betts-Henderson J, Cain K, et al. Loss of thymidine kinase 2 alters neuronal bioenergetics and leads to neurodegeneration. *Human molecular genetics* 2010; **19**(9): 1669-77.

33. Wan J, Masuda T, Hackler L, Jr., et al. Dynamic usage of alternative splicing exons during mouse retina development. *Nucleic acids research* 2011; **39**(18): 7920-30.

34. McNeely KC, Dwyer ND. Cytokinesis and postabscission midbody remnants are regulated during mammalian brain development. *Proceedings of the National Academy of Sciences of the United States of America* 2020; **117**(17): 9584-93.

35. Carvunis AR, Rolland T, Wapinski I, et al. Proto-genes and de novo gene birth. *Nature* 2012; **487**(7407): 370-4.

36. Morris DR, Geballe AP. Upstream open reading frames as regulators of mRNA translation. *Molecular and cellular biology* 2000; **20**(23): 8635-42.

37. Chen HH, Tarn WY. uORF-mediated translational control: recently elucidated mechanisms and implications in cancer. *RNA biology* 2019; **16**(10): 1327-38.

38. Wu Q, Wright M, Gogol MM, Bradford WD, Zhang N, Bazzini AA. Translation of small downstream ORFs enhances translation of canonical main open reading frames. *The EMBO journal* 2020; **39**(17): e104763.

39. Mackowiak SD, Zauber H, Bielow C, et al. Extensive identification and analysis of conserved small ORFs in animals. *Genome biology* 2015; **16**: 179.

40. Crappé J, Van Criekinge W, Trooskens G, et al. Combining in silico prediction and

ribosome profiling in a genome-wide search for novel putatively coding sORFs. *BMC genomics* 2013; **14**: 648.

41. Schwinn MK, Machleidt T, Zimmerman K, et al. CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide. *ACS chemical biology* 2018; **13**(2): 467-74.

42. Diallo LH, Tatin F, David F, et al. How are circRNAs translated by non-canonical initiation mechanisms? *Biochimie* 2019; **164**: 45-52.

43. Yang Y, Fan X, Mao M, et al. Extensive translation of circular RNAs driven by N(6)-methyladenosine. *Cell research* 2017; **27**(5): 626-41.

44. Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. The biogenesis, biology and characterization of circular RNAs. *Nature reviews Genetics* 2019; **20**(11): 675-91.

45. Wethmar K, Barbosa-Silva A, Andrade-Navarro MA, Leutz A. uORFdb--a comprehensive literature database on eukaryotic uORF biology. *Nucleic acids research* 2014; **42**(Database issue): D60-7.

46. Zhang H, Wang Y, Lu J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends in biochemical sciences* 2019; **44**(9): 782-94.

47. Somers J, Pöyry T, Willis AE. A perspective on mammalian upstream open reading frame function. *The international journal of biochemistry & cell biology* 2013; **45**(8): 1690-700.

48. Young SK, Wek RC. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *The Journal of biological chemistry* 2016; **291**(33): 16927-35.

49. Rapicavoli NA, Poth EM, Blackshaw S. The long noncoding RNA RNCR2 directs mouse retinal cell specification. *BMC developmental biology* 2010; **10**: 49.

50. Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *Journal of cell science* 2007; **120**(Pt 15): 2498-506.

51. Karali M, Banfi S. Non-coding RNAs in retinal development and function. *Human genetics* 2019; **138**(8-9): 957-71.

52. Liang WC, Wong CW, Liang PP, et al. Translation of the circular RNA circβ-catenin

promotes liver cancer cell growth through activation of the Wnt pathway. *Genome biology* 2019; **20**(1): 84.

53. Liu Y, Li Z, Zhang M, et al. Rolling-translated EGFR Variants Sustain EGFR Signaling and Promote Glioblastoma Tumorigenicity. *Neuro-oncology* 2020.

54. Wang S, Mao C, Liu S. Peptides encoded by noncoding genes: challenges and perspectives. *Signal transduction and targeted therapy* 2019; **4**: 57.

55. Makarewich CA, Olson EN. Mining for Micropeptides. *Trends in cell biology* 2017; **27**(9): 685-96.

56. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 2011; **17**(1): 10-2.

57. Joshi N, Fass J. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. 2011.

58. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 2009; **10**(3): R25.

59. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic acids research* 2019; **47**(8): e47.

60. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 2014; **15**(12): 550.

61. Zhang P, He D, Xu Y, et al. Genome-wide identification and differential analysis of translational initiation. *Nature communications* 2017; **8**(1): 1749.

62. Conesa A, Nueda MJ, Ferrer A, Talón M. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics (Oxford, England)* 2006; **22**(9): 1096-102.

63. Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in biosciences = Theorie in den Biowissenschaften* 2012; **131**(4): 281-5.

64. Chothani S, Adami E, Ouyang JF, et al. deltaTE: Detection of Translationally Regulated Genes by Integrative Analysis of Ribo-seq and RNA-seq Data. *Current protocols in molecular*

biology 2019; **129**(1): e108.

65. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* 2012; **16**(5): 284-7.

66. Kassambara A, Mundt F. Package 'factoextra'. *Extract and visualize the results of multivariate data analyses* 2017; **76**.

67. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:13033997* 2013.

68. Gao Y, Zhang J, Zhao F. Circular RNA identification based on multiple seed matching. *Briefings in bioinformatics* 2018; **19**(5): 803-10.

69. Zhang XO, Dong R, Zhang Y, et al. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome research* 2016; **26**(9): 1277-87.

70. Zhang J, Chen S, Yang J, Zhao F. Accurate quantification of circular RNAs identifies extensive circular isoform switching events. *Nature communications* 2020; **11**(1): 90.

71. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* 2013; **14**(4): R36.

72. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 2008; **26**(12): 1367-72.

73. Zdobnov EM, Apweiler R. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics (Oxford, England)* 2001; **17**(9): 847-8.

Figures and Tables

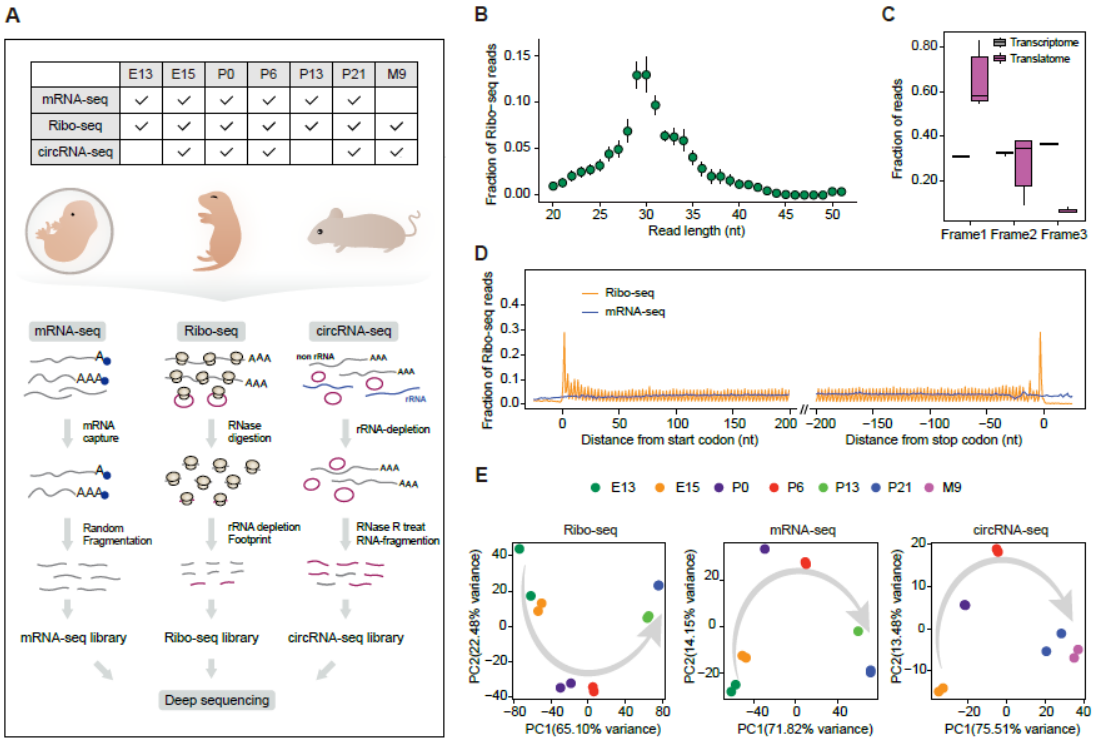

Figure 1. Overview of gene expression of the developing mouse retina. (A) A schematic illustration of the experimental design. (B) Length distribution of RPFs that mapped to CDSs of protein-coding genes, with a peak at 29-30 nt. (C) Frame distribution of Ribo-seq and mRNA-seq reads among all the CDSs, showing a clear frame preference for Ribo-seq reads and a uniform frame distribution for mRNA-seq reads. (D) Metagene analysis of read distribution around the start and stop codons of the CDSs, showing a 3-nt periodicity of Ribo-seq reads. (E) Principal component analysis of Ribo-seq, mRNA-seq and circRNA-seq data sets, respectively.

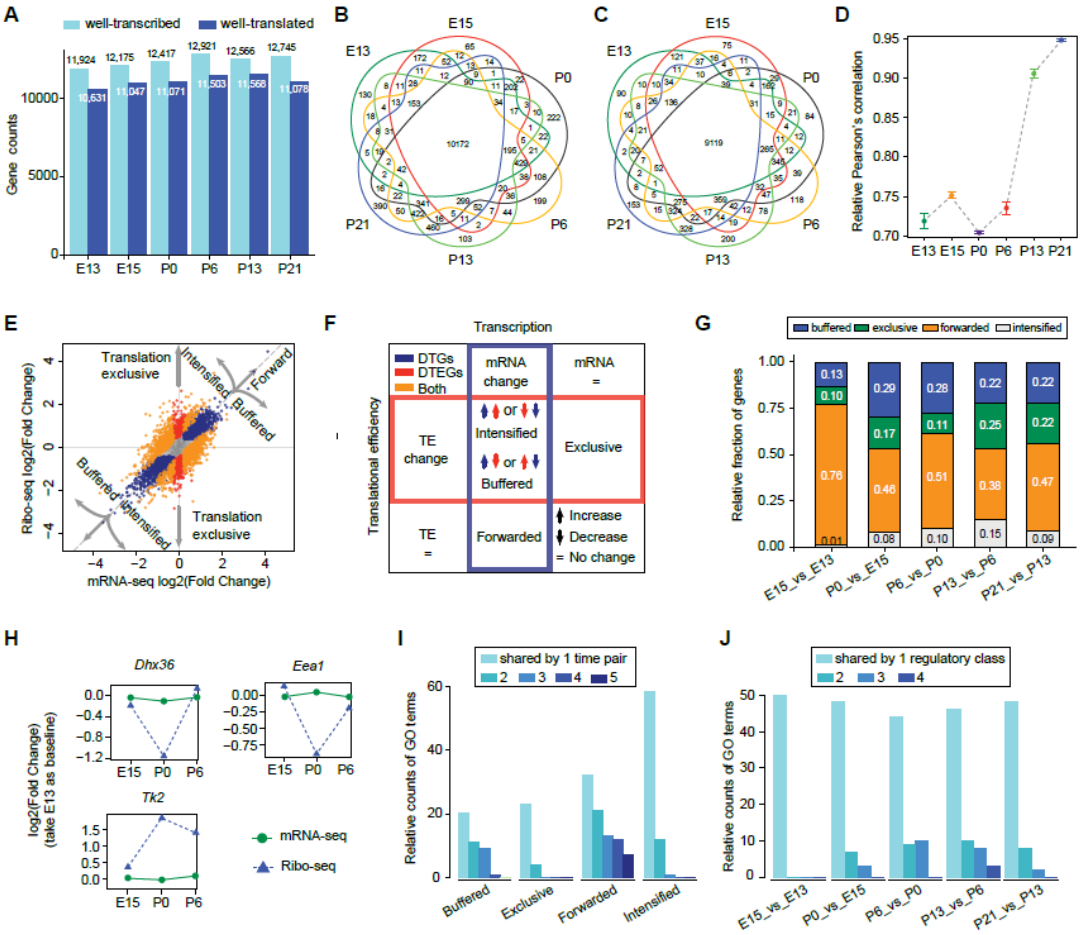

Figure 2. Transcriptional and translational characterization. (A) Number of well-transcribed and well-translated protein-coding genes in each stage. (B) Overlap of well-transcribed and (C) well-translated genes across different developmental stages. (D) Correlation of expression levels throughout development between transcriptomes and translatomes. (E) Log-fold changes in the mRNA and ribosome occupancy, taking E15 versus P0 as an example. (F) The interplay between DTGs and DTEGs showing categories of gene expression regulation. (G) Relative fractions of four categories of genes in each pairwise comparison between adjacent stages. (H) Ribosome occupancy and mRNA changes of three well-studied genes: Dhx36, Eea1 and Tk2 between E13 and P6. Lines display fold changes at Ribo-seq (yellow) and mRNA-seq (grey) level based on E13. (I) Overlap of enriched GO terms for the same category of genes across all five comparisons of adjacent stages. (J) Overlap of enriched GO terms for different categories of genes in each adjacent stage comparison. Colors represent degrees of term overlap among the four regulatory types of gene expression: buffered, exclusive, forwarded and intensified.

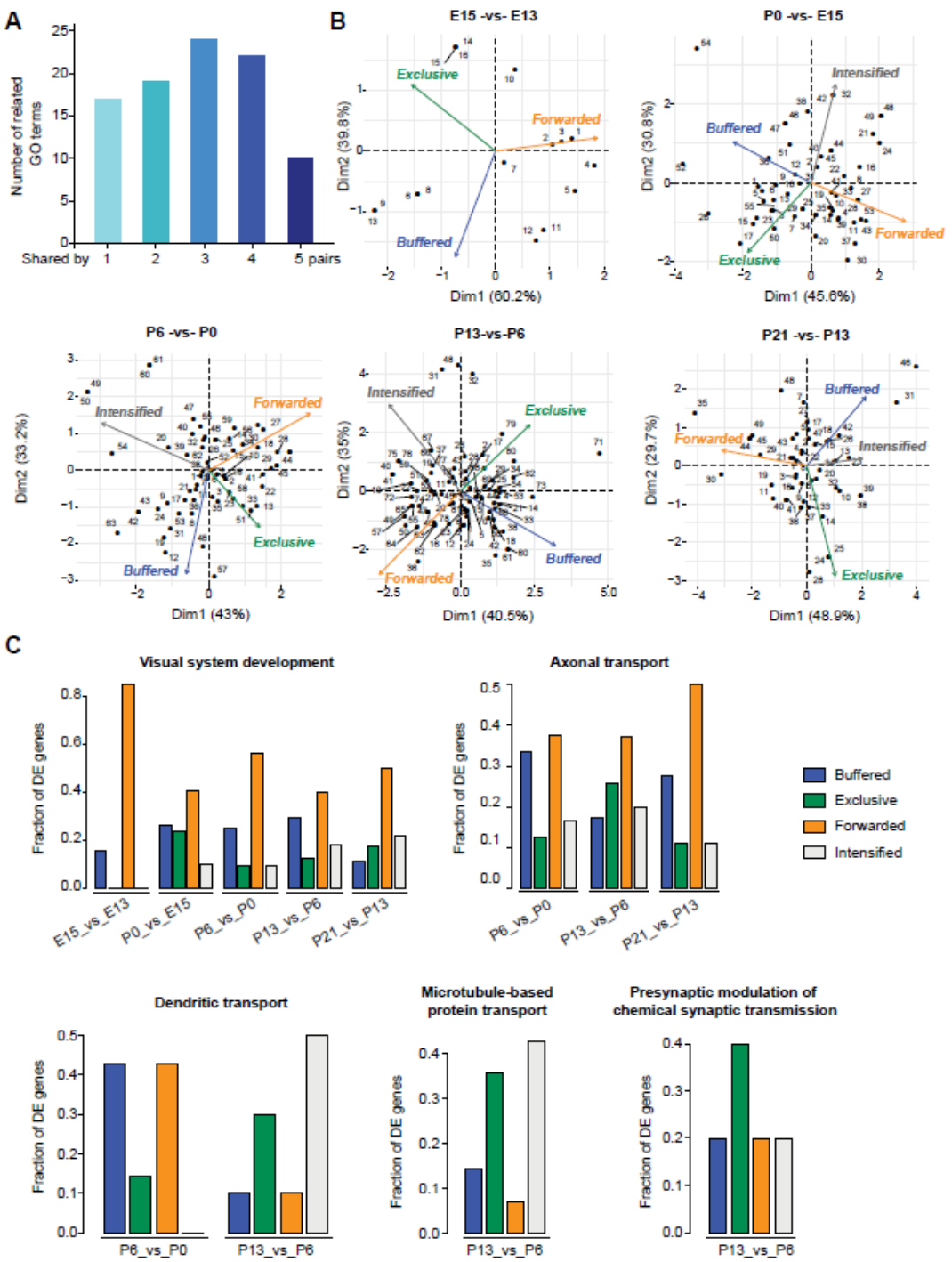

Figure 3. Dissecting Translational regulation in the developing mouse retina. (A) Frequency distribution of occurrences of enriched GO terms in all comparisons of adjacent stages. (B) Scatter plot of PCA of each developmental stage separating the manifestations of individual GO terms within the global regulatory programs (Methods). Each numbered point represents a functional GO term, and its position along each axis indicates the relative contribution of transcriptional and translational regulation to the overall differential patterns. For each

arrangement, the assigned GO term can be found in Table S4. (C) Examples of stage-specific coregulatory functional arrangements. The barplots show the relative fractions of expressed genes in retina with distinct regulatory modes.

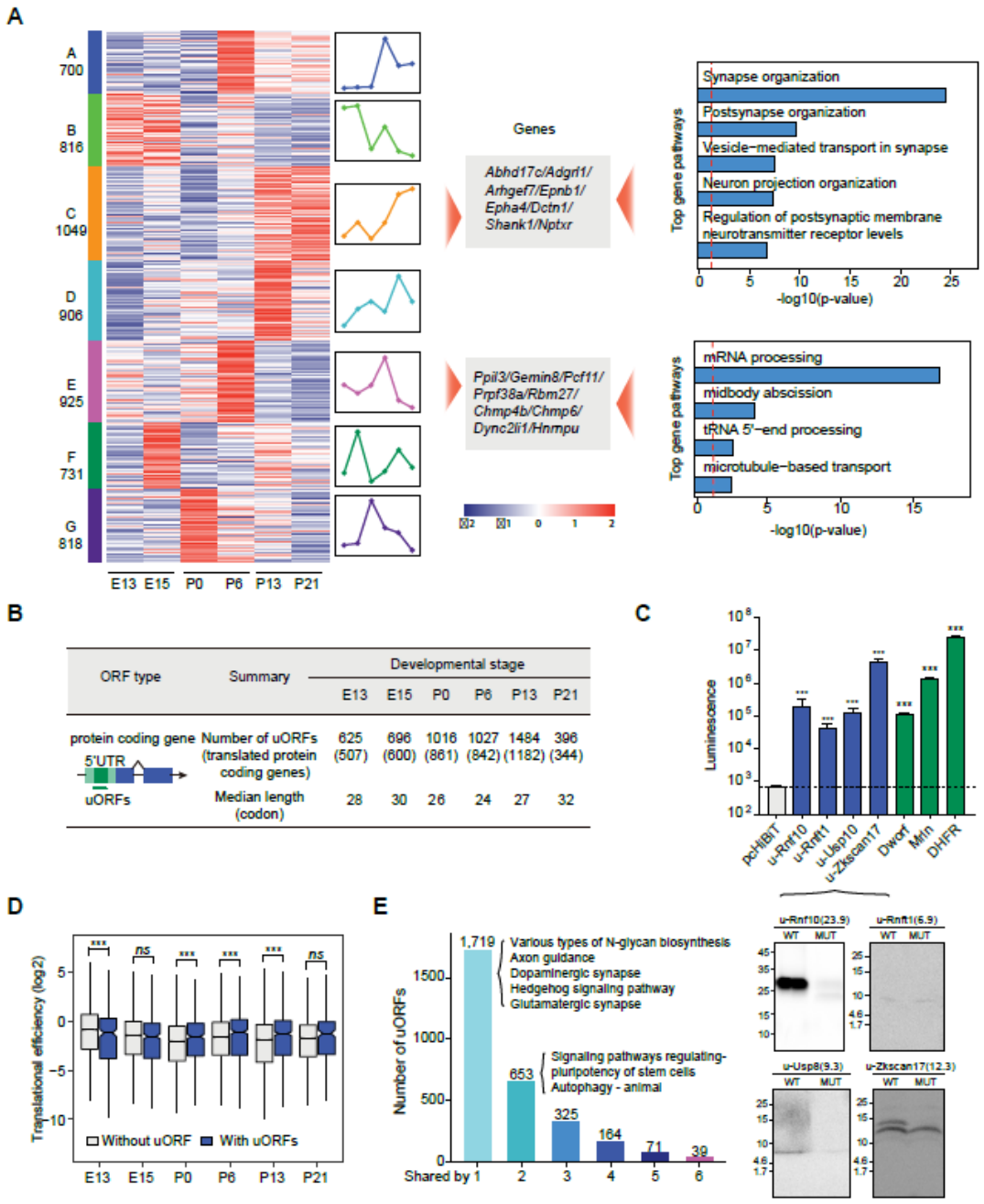

Figure 4. Analysis of translational efficiencies across retinal development. (A) Hierarchically-clustered heatmap of TE of 5,945 protein-coding genes (FDR < 0.05). Different colors indicate different clusters; numbers indicate significant DTE genes in the cluster; and temporal patterns of TE changes are portrayed through lines. Representative genes and enriched GO terms in the two largest clusters are shown on the right. (B) Number of actively translated uORFs detected in each stage by using Ribo-TISH. The uORF distribution is shown below. (C) Translation validation of uORFs: Luminescence of uORFs in IVT assays were

shown in the left panel. Western blots of uORF coded peptides (HiBiT tagged) and corresponding ATG mutants were shown in the right panel, where pcHiBiT is a negative control; Dworf and Mrln are positive control of lncRNAs; and DHFR is a protein-coding control. All candidates were compared with the negative control and the dashed line indicates the luminescence of the negative control. Note: ns, $P > 0.05$; ***, $P < 0.001$. (D) Comparison of TEs between protein-coding genes with and without uORFs in each stage. (E) Frequency distribution of occurrences of uORFs during retinal development, showing developmental stage-specific usage. Colors represent numbers of uORFs with different usage frequency during the mouse retinal development.

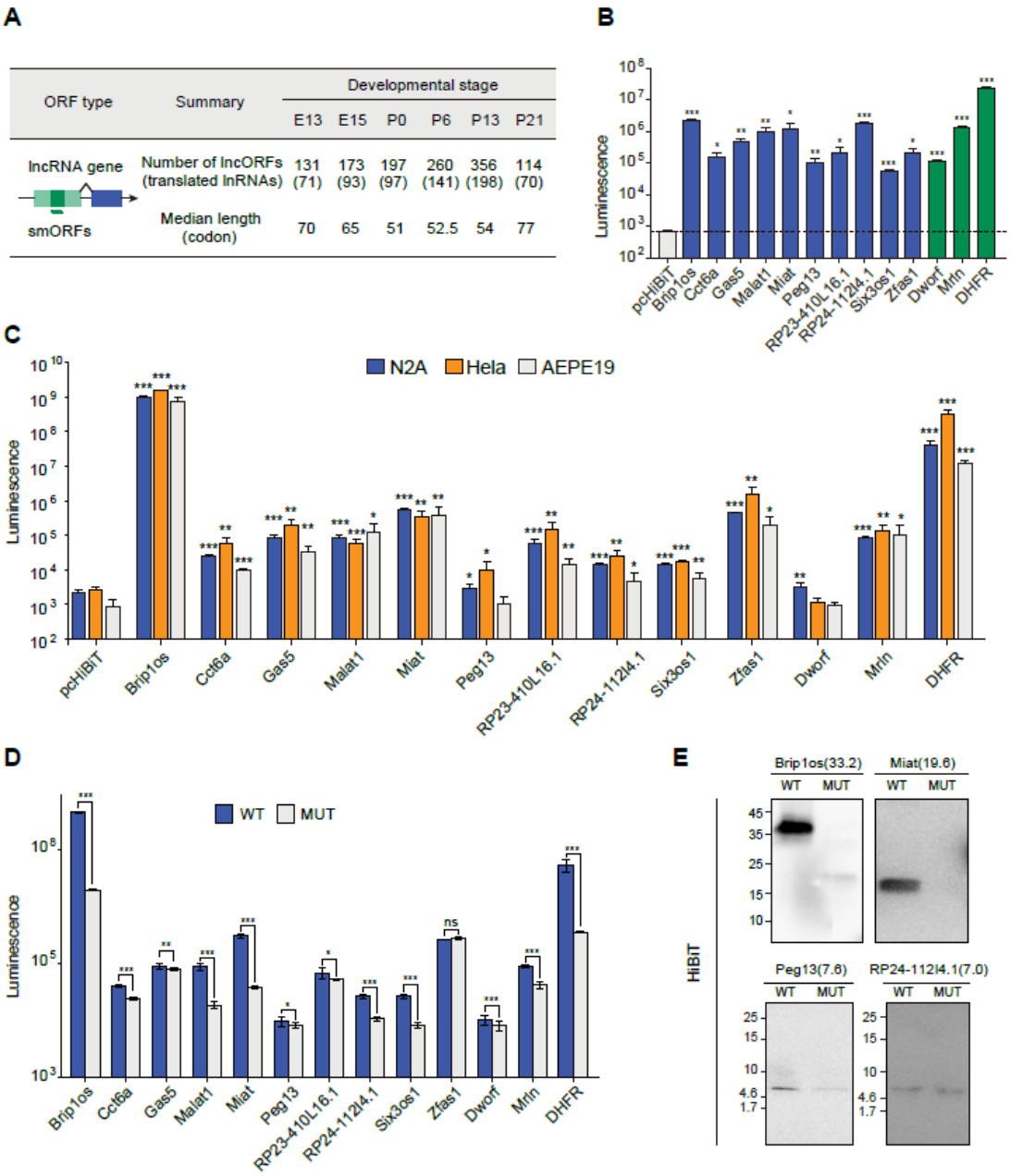

Figure 5. Translation of lncRNAs. (A) Number of actively translated lncORFs detected in each stage. (B) Luminescence of 10 candidate lncORFs in IVT assays. pcHiBiT is a negative control; *Dworf* and *Mrln* are positive controls; and *DHFR* is a protein-coding control. All candidates are compared with the negative control and the dashed line indicates the luminescence of the negative control. Note: ns, $P > 0.05$; ***, $P < 0.001$. (C) Luminescence of 10 candidate lncORFs in N2A, Hela and ARPE19 cells. (D) Luminescence of all lncORFs and corresponding ATG mutant ORFs in N2A cells. (E) Western blots of lncORF coded peptides (HiBiT tagged) and corresponding ATG mutants. Note: ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

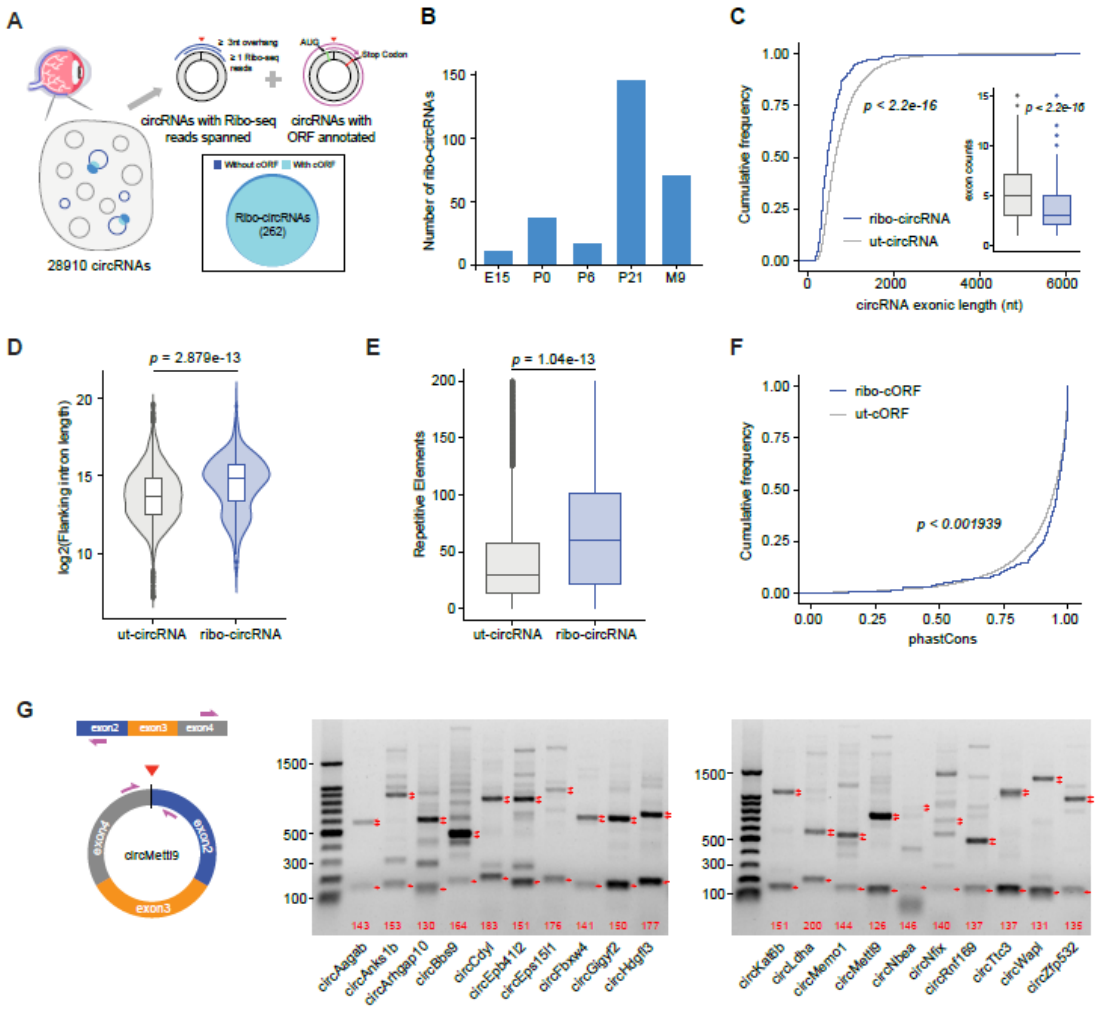

Figure 6. Translation of circRNAs. (A) Schematic overview of ribosome associated-circRNA detection. (B) Number of ribo-circRNAs identified in each stage. (C) Cumulative plot of exonic length for untranslated circRNAs (ut-circRNAs, black) and ribo-circRNAs (green). The inserted box plot showing exon number differences between ut-circRNAs and Ribo-circRNAs. (D) Violin plot of flanking intron length for ribo-circRNAs and ut-circRNAs. (E) Box plot comparing the repetitive elements in the flanking introns between ribo-circRNAs and ut-circRNAs. (F) Cumulative plot of sequence conservation (phastCons score) for ribo-cORFs (green) and ut-cORFs (black). (G) Schematic overview of circRNA divergent primer design (left) and experimental validation of the rolling circle cDNA products from circRNAs (right); circLmnb1(novel).