Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

Cell Rep. 2016 November 22; 17(9): 2460–2473. doi:10.1016/j.celrep.2016.10.074.

NRL-Regulated Transcriptome Dynamics of Developing Rod Photoreceptors

Jung-Woong Kim[1,2,7], **Hyun-Jin Yang**[1,7], **Matthew John Brooks**[1,7], **Lina Zelinger**[1], **Gökhan Karakülah**[1,*], **Norimoto Gotoh**[1,3], **Alexis Boleda**[1], **Linn Gieser**[1], **Felipe Giuste**[1], **Dustin Thad Whitaker**[1,4], **Ashley Walton**[1], **Rafael Villasmil**[5], **Jennifer J. Barb**[6], **Peter Jonathan Munson**[6], **Koray Dogan Kaya**[1], **Vijender Chaitankar**[1], **Tiziana Cogliati**[1], and **Anand Swaroop**[1,$]

[1]Neurobiology, Neurodegeneration & Repair Laboratory, National Institutes of Health, Bethesda, MD 20892, USA

[2]Department of Life Science, Chung-Ang University, Seoul 06974, Republic of Korea

[3]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

[4]Texas A&M Institute for Neuroscience, Texas A&M University, College Station, TX 77843, USA

[5]Flow Cytometry Core, National Eye Institute (NEI), National Institutes of Health, Bethesda, MD 20892, USA

[6]Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA

SUMMARY

Gene regulatory networks (GRNs) guiding differentiation of cell types and cell assemblies in the nervous system are poorly understood because of inherent complexities and interdependence of signaling pathways. Here, we report transcriptome dynamics of differentiating rod photoreceptors in the mammalian retina. As the transcription factor NRL determines rod cell fate, we performed expression profiling of developing NRL-positive (rods) and NRL-negative (S-cone like) mouse photoreceptors. We identified a large-scale, sharp transition in the transcriptome landscape between postnatal day 6 and 10 concordant with rod morphogenesis. Rod-specific temporal DNA methylation corroborated gene expression patterns. *De novo* assembly and alternative splicing

[$]To whom correspondence should be addressed: Anand Swaroop, Ph.D., N-NRL, National Eye Institute, 6 Center Drive, MSC0610, Bethesda, MD 20892. Phone: 301-435 5754; swaroopa@nei.nih.gov, Lead Contact: Dr. Anand Swaroop at swaroopa@nei.nih.gov.
[*]Current address: Dokuz Eylul University, Izmir International Biomedicine and Genome Institute, Inciralti, 35340, Turkey
[7]Additional footnotes: Co-first authors

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Contributions

Overall Conceptualization, J.K., H.Y., N.G. and A.S.; Methodology and Investigation, J.K., H.Y., N.G., L.G., D.T.W., R.V.; Transcriptome Data Analysis, M.J.B., V.C., J.J.B., A.W., A.B., and P.M.; Data Curation, M.J.B.; Network Analysis, V.C., F.G.; De novo transcript assembly and alternate splicing analyses, G.K.; Database, K.D.K., M.J.B.; Writing – Original Draft, H.Y., J.K., M.J.B, L.Z., A.S.; Writing – Review & Editing, J.K., H.Y., M.J.B., L.Z., G.K., V.C., T.C., and A.S.; Funding Acquisition, A.S.; Supervision and Project Administration, A.S.

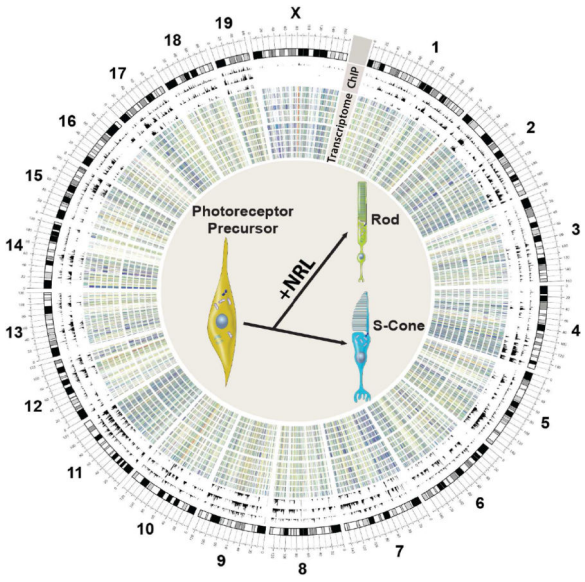
Accession Numbers

RNA-seq and microarray data are deposited in Gene Expression Omnibus (GSE74660, GSE74657).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

analyses revealed previously un-annotated rod-enriched transcripts and the role of NRL in transcript maturation. Furthermore, we defined the relationship of NRL with other transcriptional regulators and downstream cognate effectors. Our studies provide the framework for comprehensive system-level analysis of the GRN underlying the development of a single sensory neuron, the rod photoreceptor.

Graphical abstract



INTRODUCTION

Diversity in neuronal morphology and function is produced by coordinated and controlled changes in gene expression patterns that are established by gene regulatory networks (GRNs), which provide roadmaps and topological information of regulatory and effector genes guiding specific biological processes (Peter and Davidson, 2011). Combinatorial actions of select transcriptional regulators on cis-regulatory genomic modules (Guillemot, 2007; Spitz and Furlong, 2012) generate context, space, and time dependent gene expression states (Allan and Thor, 2015; Tsankov et al., 2015). The developmental transcriptome is dynamic and yields a deterministic framework for downstream physiological functions. Thus, comprehensive mapping of the transcriptomic landscape of a specific neuronal cell type can yield better understanding of mechanisms underlying the acquisition and maintenance of its unique identity.

Retinal photoreceptors are sensory neurons that capture light and initiate the visual process (Lamb et al., 2007). Rod and cone photoreceptors have unique morphology to maximize their functional output, with numerous membrane discs in outer segments for photon capture and specialized ribbon synapses for signal transmission. Cones contain visual opsins with distinct absorption spectra and provide high acuity daylight color vision, whereas rods mediate low-light vision using a single visual pigment, rhodopsin. Rod photoreceptors dominate the retina of most mammals (70–80% of all retinal cells), including mice and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

humans, and are especially vulnerable to genetic changes in retinal and macular diseases (Wright et al., 2010).

Differentiation of rod photoreceptors spans a long temporal window far beyond their birth (approximately three weeks in mice), thereby offering an attractive model to investigate transcriptome dynamics and GRN that control their distinctive neuronal morphology and function. Several transcription factors (TFs), including OTX2, RORB and PRDM1, regulate the photoreceptor lineage in developing retina (Brzezinski et al., 2013; Jia et al., 2009; Katoh et al., 2010; Nishida et al., 2003; Roger et al., 2014); however, rod fate is critically dependent on expression and activity of the Maf-family leucine zipper TF, NRL (Swaroop et al., 2010). Photoreceptor precursors fated to be rods acquire characteristics of short-wave length (S) cones in the absence of *Nrl* (*Nrl*$^{-/-}$) (Mears et al., 2001), whereas expression of *Nrl* in developing cones leads to rod differentiation (Oh et al., 2007). NRL collaborates with cone-rod homeobox CRX to activate rhodopsin and other rod-specific genes (Chen et al., 1997; Mitton et al., 2000). A key downstream target of NRL, the orphan nuclear receptor NR2E3, actively suppresses cone genes and consolidates the rod cell fate (Cheng et al., 2006; Oh et al., 2008). In addition, estrogen-related receptor ESRRB, myocyte enhancer factor MEF2C, histone lysine-specific demethylase KDM5B and transcription-splicing protein NONO are among NRL targets that regulate specific aspects of rod development and survival (Hao et al., 2012; Hao et al., 2011; Onishi et al., 2010; Yadav et al., 2014). Thus, genome-wide targetome studies of NRL, CRX (Corbo et al., 2010; Hao et al., 2012), and other downstream TFs can be integrated with transcriptome profiles to assemble GRN modules associated with rod differentiation (Yang et al., 2015).

Here, we report directional RNA-seq analysis of developing mouse rod photoreceptors and of *Nrl*$^{-/-}$ photoreceptors fated to be rods, thereby directly determining the impact of perturbing a key hub in the rod-specific GRN. We tracked the dynamic changes occurring in the NRL-centered GRN and their impact on rod morphogenesis by integrating transcriptome, targetome and DNA methylome datasets. We further define previously un-annotated NRL-regulated transcripts by *de novo* assembly and assess the relevance of alternative splicing during photoreceptor development. Our studies thus provide a platform for system-level analysis of a single type of sensory neuron, the rod photoreceptor.

RESULTS

Transcriptome Profiling of Developing Rod Photoreceptors

In mice, the birth (final mitosis) of rod photoreceptors overlaps with the generation of all other retinal cell types and begins as early as embryonic day (E)12 (Figure 1A); however, a majority of rods are generated at postnatal day (P)0 to P2 (Akimoto et al., 2006; Carter-Dawson and LaVail, 1979). Though P2 mouse retina is enriched for newborn rods, robust up-regulation of rhodopsin (encoded by *Rho*) and other phototransduction genes is detected later at P4. Outer segment formation and synaptogenesis begin at or after P6 (Blanks et al., 1974; Obata and Usukura, 1992) and continue through P10-P12. After eye opening at P14, synaptic changes are predicted based on activity until rods reach functional maturity by P21 (Figure 1A). To elucidate temporal dynamics of gene regulation, we generated rod photoreceptor transcriptome throughout development and then integrated differential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

expression analysis with NRL ChIP-seq data (Hao et al., 2012), and transcription factor binding motifs (Figure 1B).

Transcriptome profiling was performed using exon microarrays and directional RNA-seq platforms. We flow-sorted rods from *Nrl*p-GFP mouse retina (Akimoto et al., 2006) at P2, P4, P6, P10, P14 and P28 for RNA-seq and P2, P4, P6, P8, P10, P12, P21 for microarray analysis (Figure 1A, B; Figure S1). The two datasets demonstrated concordant transcriptome dynamics during differentiation. Principal component analysis (PCA) revealed that developmental stage accounted for the largest variance among datasets in both platforms (Figure 1C). Biological replicates of each time point clustered together (Figure 1C), confirming high reproducibility. The rod transcriptome data were separated in an ordered fashion along PC1 and suggested a linear progression of gene expression changes during development (Figure 1C). The correlation between P21 microarray and P28 RNA-seq data (Figure 1D) indicated similar trends in the two platforms. Nonetheless, directional RNA-seq displayed a greater dynamic range, higher sensitivity in detecting expression changes [7,841 differentially expressed (DE) genes for P28/P2 RNA-seq having ≥2 fold change, ≤0.01 FDR, and ≥1 FPKM at any time point versus 2,533 DE genes with ≥2 fold change and ≤0.05 FDR for P21/P2 microarray], and a negligible background noise compared to the hybridization-based microarray methodology. We, therefore, used directional RNA-seq data for subsequent comparative and integrated analysis. To facilitate rapid access, analysis and visualization of datasets, we also developed an integrative platform, Retseq (https://retseq.nei.nih.gov).

Transition of Transcriptome Landscape in Developing Mouse Rods between P6 and P10

Hierarchical clustering and unsupervised Pearson's correlation analysis of averaged, normalized Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values at individual time points showed a decisive shift in rod transcriptome as a whole, and in specific transcripts with FPKM >5 in particular, between P6 and P10 (Figure 2A). A similar shift was apparent between P6 and P8 on PC1 analysis of microarray data (Figure 1C). Our data suggest that this transition in transcriptome landscape is highly synchronized and regulated, occurring within a relatively short time window, and constitutes a critical step in rod differentiation.

Specificity of the transcriptome data was validated by expression of known rod genes at P28 (Figure 2B). Furthermore, while a subset of cone genes were transiently expressed in rods at early time points, their expression was minimal or undetected by P28 (Figure 2B) (Kim et al., 2016). Pan-photoreceptor and pan-neuronal genes were highly expressed, whereas genes specific for inner retinal neurons (i.e., bipolar, ganglion, horizontal, and amacrine cells) and Müller glia were barely detectable in mature rods (Figure 2B).

Redundancy-removed gene ontology (GO) analysis (http://david.abcc.ncifcrf.gov) of the top 1,000 highly expressed genes at each time point revealed both constitutive and stage-specific biological processes across development (data not shown). For example, early expressed genes were associated with RNA processing and metabolism, whereas photoreceptor development was among the highly enriched pathways at later stages. Enrichment of genes for ATP synthesis coupled proton transport and hexose catabolic process was consistent with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

high-energy requirement for photoreceptor morphogenesis and function (Kooragayala et al., 2015).

Transcriptome of Developing S-cone like *Nrl*-KO Photoreceptors

To investigate the precise role of NRL in mediating transcriptome dynamics, we performed directional RNA-seq analyses of *Nrl*p-GFP;*Nrl*$^{-/-}$ (or referred as *Nrl*-KO hereafter) photoreceptors (Akimoto et al., 2006). The data validated complete absence of rod-specific phototransduction genes and high expression of cone genes, as shown earlier by microarrays (Akimoto et al., 2006; Yoshida et al., 2004). PCA analysis revealed a major divergence in transcriptome of S-cone like *Nrl*-KO photoreceptors compared with wild type rods (Figure 3A). The transcriptome of the two genotypes clustered at P2, but the difference increased dramatically as differentiation progressed (Figure 3A). Differential expression (DE) analysis identified 4,585 protein-coding transcripts (3,888 genes) showing significant up- or down-regulation in *Nrl*-KO photoreceptors (FPKM >5 in at least one time point; fold change >2; FDR < 1%; Figure 3B). As anticipated from the PCA analysis, the number of DE transcripts increased with rod morphogenesis and maturation (from 113 DE transcripts at P2 to 3962 DE transcripts at P28). The dynamic transition in the transcriptome landscape from P6 to P10 that was apparent in rods was barely detectable in *Nrl*-KO photoreceptors (Figure 3B).

To discriminate between direct and indirect targets of NRL, we integrated DE transcripts between rods and *Nrl*-KO photoreceptors with the NRL ChIP-seq data (Hao et al., 2012). As predicted, a subset of DE transcripts carried NRL ChIP-seq peaks (Figure 3C). Of these, 40 different genes encoded TFs that are putative direct targets of NRL (Figure 3D) and represent a starting point for potential secondary regulation functions. As an example, *Tnfaip3*, previously implicated in negative regulation of cell death pathway, included two NRL binding sites and showed dramatically reduced expression in *Nrl*-KO photoreceptors (Figure 3D), suggesting its role in functional maintenance and/or survival of rods.

Stage-specific DNA Methylation in Photoreceptor Genes

Rod-specific phototransduction genes, such as *Rho*, *Gnat1*, and *Cnga1*, exhibited a dramatic up-regulation during P6 to P10, and a subset of cone-specific genes were detectable in immature rods until a later stage (see Figure 2B). To correlate DNA methylation to changes in global gene expression (Suzuki and Bird, 2008), we performed reduced representation bisulfite sequencing (RRBS) (Gu et al., 2011) using genomic DNA from P2, P10, and P28 rods. As predicted, the level of DNA methylation both in the gene promoter area and in the gene body showed a negative correlation with gene expression at all three stages (Figure S2). A subset of rod-specific genes (e.g., *Rho* and *Pde6b*) showed high level of DNA methylation at their respective promoter at P2; however, by P28, the methylation was undetectable (Figure 4A). Interestingly, a greater number of rod genes had high level of DNA methylation in the gene body than in the promoter (Figure 4A). The degree of DNA methylation in the gene body of rod-specific genes was greatly reduced by P10 to P28, consistent with delayed onset of their expression. As expected, *Nrl* exhibited constantly low DNA methylation at both promoter and gene body in rods at all times (Figure 4A). We further noted prominent DNA methylation in the gene body of cone-specific genes consistent with their lack of expression, while the promoter region of cone genes remained unmethylated (Figure 4A).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Almost all non-photoreceptor retinal genes were unmethylated in their promoter region, though most genes acquired medium-to-high DNA methylation in the gene body as rod maturation progressed (Figure 4B). No DNA methylation was detected in the promoter region of housekeeping genes, consistent with their constitutively high expression, with a few having medium level of DNA methylation in the gene body (Figure 4B).

De Novo Assembly and Identification of Previously Un-annotated Transcripts

Our time series of developing mouse rods and S-cone like *Nrl*-KO photoreceptors contained approximately 80,000 transcripts that were not annotated in the Ensembl database. Of these, over 20,700 constituted isoforms of known genes and 2,840 were multi-exonic intergenic transcripts (NIMETs) with FPKM value ≥1 in all replicates of at least one time point (Figure 5A; Figure S3). Mono-exonic transcripts and those expressed at <1 FPKM were excluded from further analysis since these may represent transcriptional noise. Of 2,840 NIMETs, almost 15% showed <1 FPKM in RNA-seq data from other tissues examined (Consortium, 2004). Using four criteria (intergenic, multi-exonic, dynamic expression pattern, and at least one conserved domain), we identified 222 potentially protein-coding (NPC) NIMETs, containing 775 putative conserved domains (E-val <0.05). Notably, a sharp transition in expression between P6 and P10 rods was also evident in these transcripts (Figure 5A, left panel). However, the transition seemed more gradual, delayed or even absent in *Nrl*-KO photoreceptors (Figure 5A, right panel). We identified three major clusters based on expression patterns (Figure 5A). Members of cluster 1 showed a gradual increase in expression during rod differentiation, and this increase was not detected in the absence of NRL (Figure 5A), as is the case with rod-specific genes. A subset of transcripts in cluster 1 exhibited delayed onset of expression, suggesting their function in rod maturation. In contrast, transcripts in clusters 2 and 3 were highly expressed in *Nrl*-KO photoreceptors compared to rods (Figure 5A) and likely represented cone-enriched genes.

Conserved functional domains of NPC transcripts were detected with HMMER against Pfam database (Finn et al., 2010). Identified domains spanned various functions from RNA processing (e.g., *M*ed19, PF10278) to structural domains (e.g., Golgin_A5, PF09787) and protein degradation (e.g., Peptidase_S49_N, PF08496). Top represented domains (identified at least 10 times) were grouped into families based on their putative function annotated in Pfam database. The most frequently occurring domains were of unknown function, but the second largest family was related to transcription regulation and included several zinc finger domains (Figure 5B). Other highly represented families included PDZ and FERM domains involved in signaling and protein localization (Figure 5B). qRT-PCR analysis validated mature rod-specific expression of select NPC transcripts (Figure 5C), arguing in favor of their functional relevance.

We superimposed genomic locations of NPC transcripts with NRL and CRX ChIP-seq peaks using 1 kb as maximal allowed distance between the peak and the transcript start or end point. Of the 222 NPC transcripts, 36 were potentially regulated by both NRL and CRX, while NRL alone regulated three and CRX 20 additional NPCs. Evolutionary conserved regions overlapped with predicted gene structure of 67% NPCs carrying both NRL and CRX ChIP-seq peaks (24 out of 36), with additional 10 NPCs showing a partial overlap. For

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

example, one NPC transcript, TCONS_00123129 (part of cluster 1), contained three significant Pfam domains (L27_1 - PF09058, hDGE_amylase - PF14701, and CD36 - PF01130) with strong evolutionary conservation in the gene structure, and its promoter region included NRL and CRX ChIP-seq peaks (Figure 5D).

As a proof of principle, we focused on TCONS_00068375, which is a part of potentially rod-enriched cluster 1 (Figure 5A; Figure S3) and contains NRL and CRX ChIP-seq peaks. This transcript contains a G-patch domain (PF01585), present in RNA binding proteins, and is located between two annotated genes *Slc22a28* and *Slc22a29* that show very low or no expression in rods (Figure S4A). qRT-PCR and *in situ* hybridization (ISH) validated RNA-seq data and revealed its up-regulation at P28 (Figure S4B, S4C); this increase was not observed in *Nrl*-KO retina (Figure S4B). *In vivo* knock down of TCONS_00068375 transcript in developing mouse retina resulted in thinning of outer nuclear layer (ONL) (Figure S4D), suggesting its potential function in rod photoreceptors.

ISH of two additional NPCs recapitulated their dynamic expression pattern in the ONL (Figure 5E). While TCONS_00044530 was undetectable at P6 but observed at P28, TCONS_00069915 transcripts were prominent at P6 but absent at P28 (Figure 5E). TCONS_00044530 transcript also belongs to cluster 1 (Figure 5A; Figure S3) and is predicted to encode a protein with nuclear RNA-splicing-associated domain (SR-25, PF10500), whereas TCONS_00069915 (cluster 3) likely encodes endonuclease/ exonuclease/ phosphatase family (Exo_endo_phos, PF03372) and syntaxin 6 (PF09177) protein domains (Figure 5A; Figure S3).

Alternative Splicing (AS) and Alternate Promoter Usage

By surveying developmental time- and NRL-dependent AS events in RNA-seq data, we identified 809 and 483 differential splicing events during the development of rods and *Nrl*-KO photoreceptors, respectively (FDR<0.01; Figure 6). A comparison of all temporal data from rods versus *Nrl*-KO photoreceptors identified 3,644 differential AS events of all five major types (FDR<0.01; Figure 6A) as well as potential alternative promoter usage and introduction of new exons. Rods and *Nrl*-KO photoreceptors revealed distinct genome-wide AS events. While transcripts with a retained intron (RI) were frequently detected in rods at P2 and P4, *Nrl*-KO photoreceptors showed increased RI and skipped exon (SE) between P6 and P10 (Figure 6A). The most dramatic shift in AS was evident in rods versus *Nrl*-KO photoreceptors from P6 to P10 (Figure 6A), similar to that detected in the broader transcriptome landscape (see Figure 2A). RI was the most common AS event in rods with highest number at P28 (Figure 6A).

The two most frequent AS events observed, SE and RI, are exemplified by two genes, clathrin light polypeptide (*Clta*) and calcium channel, voltage-dependent, L type, α1F subunit (*Cacna1f*), respectively (Figure 6B). The *Clta* gene produced three isoforms by skipping or including exons 5 and 6. In P2 rods, the two most abundant isoforms skipped either both or only exon 6; however, increased expression of exon 6-containing transcripts was observed at P28 (Figure 6B, left). *Cacna1f* transcript in mature rods at P28 did not include any intron 6 sequences, which were evident at P2; however, introns were not

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

removed in *Nrl*-KO photoreceptors, suggesting a role of NRL in transcript maturation (Figure 6B, right).

Consistent with previous observations (Hao et al., 2011; Hao et al., 2014), we identified several rod-specific alternative transcript isoforms that are regulated by NRL. Hematopoietic cell-specific Lyn substrate 1 (*Hcls1*), associated with gene regulation in the hematopoietic system (Skokowa et al., 2012), exhibited rod-specific, late onset expression of a transcript which is initiated upstream of exon 8 from a new transcription start site, suggesting alternative promoter usage (Figure 6C). Furthermore, *Hcls1* included NRL and CRX ChIP-seq peaks upstream of a retina-specific exon. ISH demonstrated *Hcls1* expression specifically in the retinal ONL at P28 (Figure 6C), implying an as yet unrecognized function of this transcription factor in rod photoreceptors.

Co-expression Network Analysis

A network of transcription factors including NRL and CRX is predicted to drive the gradual transcriptome shift during rod maturation. By employing weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008) to all transcripts expressed at ≥5 FPKM at one or more time points, we identified 11 gene modules that correlated to distinct temporal expression patterns (Figure S5). A vast majority (672 out of 803) of TF-encoding transcripts was present in modules A, B and C, which included all known rod-specific transcriptional regulators (Figure S5). Analysis of module A yielded a sub-cluster of 12 TFs including *Esrrb* and *Mef2C* (Figure S5; Rod cluster), which are known regulators of rod differentiation (Hao et al., 2012; Hao et al., 2011; Onishi et al., 2010). The ‘rod cluster’ exhibited a clear shift in expression between P6 and P10, with no or low expression in the early postnatal days and a substantially higher expression by P10 and later. This shift was not detected in *Nrl*-KO data.

NRL-centered Transcriptional Network

To construct the GRN associated with rod morphogenesis and functional maturation, we integrated differential gene expression analyses of time series and knockout samples to NRL targetome data and transcription factor binding motifs. We noted that a subset of early NRL targets were transcriptional regulators (see Figure 3D), which in turn could modulate expression of other downstream genes. We thus investigated potential secondary hubs (referred as source genes herein) that cooperate with and mediate distinct subsets of NRL function. We identified four potential secondary hubs of the NRL-centered GRN - BHLHE41, ESRRB, FOS, and NR2E3 (Figure 7A) - by differential expression analysis comparing rods and *Nrl*-KO photoreceptors during P2-P6 period and using a stringent cutoff of statistical significance and expression level (see Supplementary Methods for details). The four secondary nodes exhibited distinct expression profiles during rod differentiation and upon genetic loss of *Nrl* (Figure 7B). To construct the NRL-centered network with the four secondary hubs, potential downstream targets of source nodes were identified using the JASPAR database of DNA-binding motifs (Sandelin et al., 2004). We confined the targets of source genes to those differentially expressed in P10, P14, and P28 rods compared to P6 or earlier. The reconstructed network had 1,015 nodes (i.e., individual genes that constitute the network) and 1,414 edges (i.e., predicted interaction among any two nodes within the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

network) (Figure 7C, right). Notably, phototransduction, visual perception, eye morphogenesis, photoreceptor cell development, and non-motile primary cilium (photoreceptor outer segment) were highly represented GO terms among the nodes in NRL-centered network (Figure 7C).

DISCUSSION

Neuronal cell type-specific analysis has been limited to heterogeneous transcriptome landscapes of discrete regions of the brain (Hawrylycz et al., 2012; Tebbenkamp et al., 2014) or stem cell-derived neurons (Kaewkhaw et al., 2015; Thakurela et al., 2015; van de Leemput et al., 2014). Only a few neuron-specific transcriptomes have been reported so far (Akimoto et al., 2006; Macosko et al., 2015; Siegert et al., 2012; Telley et al., 2016; Zhang et al., 2014). Here, we used the mouse rod photoreceptor as a model sensory neuron to establish temporal transcriptome dynamics driven by the rod cell fate determination factor NRL, using two distinct platforms – RNA-seq and exon arrays – for cross-validation and deeper coverage of differentially expressed genes. We integrated multiple genome-wide expression datasets, before and after perturbing *Nrl*, and superimposed these with cis-regulatory maps obtained by NRL ChIP-seq (Hao et al., 2012), motif analysis and global DNA methylation profile. A genome level topology of “rod regulome” (Figure S6) provides a comprehensive view of the power of system-level analysis and a framework upon which additional “omics” and epigenetic data can be superimposed.

In this study, we focused on three fundamental questions: (i) are there strategic attributes in the transcriptional dynamics of developing rods; (ii) do we detect unknown or unusual elements (genes or transcripts) that complement the existing rod transcriptome landscape; and (iii) how does the NRL-centered network expand to complete rod development? The most notable characteristic we identified is the dramatic shift in transcriptome profile between P6 and P10 in developing rods. Postmitotic photoreceptors appear to retain developmental plasticity (Adler and Hatlee, 1989; Brzezinski et al., 2013; Cheng et al., 2006; Ng et al., 2011; Oh et al., 2007) until their fate is stabilized by additional changes in the differentiation program. The observed change in rod transcriptome between P6 and P10 may represent the decisive transition from early developing rods (still plastic) to functionally maturing rods (with stable cell fate). Down-regulation of *Vsx2* after P10 (see Figure 2B) is consistent with its proposed role in controlling bipolar cell genesis through inhibition of rod fate (Livne-Bar et al., 2006). Similarly, PRDM1, which regulates *Otx2* and the decision between rod and bipolar fate in some progenitors (Wang et al., 2014), is also down-regulated by P10. Low-level expression of *Vsx2* and *Prdm1* in early developing rods may indicate a molecular memory from the original retinal progenitor. Notably, we detected expression of a number of S-cone genes in developing but not in mature rods (Kim et al., 2016), consistent with the hypothesis of S-cone as a “default” fate in many (likely prenatal) photoreceptor precursors in mouse retina (Swaroop et al., 2010) and suggesting the recruitment of S-cones to rod fate to mitigate nocturnal bottleneck during early mammalian evolution (Kim et al., 2016). We believe that the P6 to P10 transition in rod transcriptome consolidates rod versus bipolar or rod versus cone alternative fates by decisively suppressing molecular footprints associated with early steps in lineage selection.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

NRL is the primary regulatory node in the rod GRN and orchestrates the consolidation of rod fate, expression of phototransduction genes, outer segment morphogenesis, assembly of pre-synapse, and functional maturation by directly or indirectly modulating the expression of most photoreceptor-enriched genes (Hao et al., 2012). Our analysis reveals previously unappreciated substructures of the NRL-centered network with four putative secondary hubs: BHLHE41 (identified by the WGCNA analysis as well), ESRRB, FOS, and NR2E3. Identification of NR2E3 provides credibility to other predicted secondary hubs. Interestingly, 1,015 genes that comprise the new network segregate into distinct clusters that are enriched in genes associated with membrane, cilia and outer segments, as well as phototransduction. These genes are differentially expressed at or after P10 and are either direct targets of NRL or are indirectly regulated through the secondary hubs. Unlike the consistently high expression of NRL, three of the secondary hubs (BHLHE41, ESRRB, and FOS) exhibit stage-dependent expression patterns, which may allow for a stepwise regulation and explain delayed onset of rod phototransduction gene expression and cilia/synapse formation. Of note, BHLHE41 is shown to participate in regulating circadian rhythm-associated genes, like *PER1*, which in turn may impact circadian-associated rod physiology and function (Green and Besharse, 2004; McMahon et al., 2014).

Control of rod development by NRL is likely fine-tuned by tertiary and higher order co-regulators and chromatin accessibility (Spitz and Furlong, 2012). Co-expression and network analyses reveal additional regulatory proteins that show a sharp increase in expression between P6 and P10, suggesting their possible role as co-regulators with NRL in controlling subsets of targets to produce rod-specific attributes. Known NRL-interacting TFs (MEF2C and ESRRB) were identified by the WGCNA analysis, together with potential regulators (HCLS1 and BHLHE41) expanding the repertoire of regulatory machinery involved in photoreceptor development.

In our study, the probability to detect true positives by genome-guided *de novo* assembly (Li et al., 2014) was augmented by temporal rod transcriptome analysis together with the expression data from $Nrl^{-/-}$ S-cone like photoreceptors. Many previously un-annotated transcripts are predicted to encode proteins with DNA-binding zinc finger Cys2-His2 (zf-C2H2) domains that might have significant regulatory functions (Najafabadi et al., 2015) in fine-tuning photoreceptor maturation. ISH and qRT-PCR validation of select transcripts and the presence of NRL and CRX ChIP-seq peaks (see Figure 5C–E) seem to corroborate our hypothesis. Indeed, knock-down of TCONS_00068375 showed thinning of the outer nuclear layer, consistent with its expression pattern and regulation by NRL. Our analysis also validated the prevalence of AS in developing photoreceptors, as implicated by whole retina studies (Liu and Zack, 2013; Wan et al., 2011). We identified a number of previously un-annotated transcript isoforms and recognized retention of intron (RI) as the dominant AS event in developing photoreceptors. Our study corroborates the reported importance of RI in differentiating mammalian cells (Wong et al., 2013; Yap et al., 2012) and further uncovers a potential role of NRL in facilitating appropriate splicing of gene products in maturing rods.

The transcriptome catalogue of rod development and associated gene regulatory network provides a reference map to investigate photoreceptor biology with potential implications for human retinal degeneration. Despite the identification of over 200 genes associated with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

photoreceptor disease (RetNet; https://sph.uth.edu/retnet/), not enough is known about the precise impact of disease-causing mutations on gene/protein networks controlling photoreceptor homeostasis. The existence of non-coding genetic variants in the human population (Genomes Project et al., 2015) and those specifically segregating in Mendelian diseases (Cartegni et al., 2002) poses a challenge in establishing genotype-phenotype relationships. The problem is even more acute in complex multifactorial diseases (e.g., age-related macular degeneration) where a vast majority of associated variants are in the non-coding region (Chakravarti et al., 2013; Fritsche et al., 2016). Previously un-annotated transcript isoforms and/or intergenic transcripts identified here may provide insights in the missing heritability.

Many complex neurodegenerative diseases afflict distinct subsets of neurons (Mattson and Magnus, 2006), emphasizing the importance of system-level assessments of individual neuronal populations. Transcriptome analysis is now possible even at a single cell level (Shapiro et al., 2013), and many genome-wide methods are being miniaturized (Buenrostro et al., 2013; Smallwood et al., 2014), making similar multidimensional approaches a possibility for even low abundance neuronal cells. We propose our integrated analysis of photoreceptor development (rods and S-cone like) as a blueprint for cell type-specific analysis of other neurons. Epigenomic landscape of adult rods and cones demonstrates unique profiles even in closely related neuronal subtypes (Mo et al., 2016). A comparison of rod and cone P28 transcriptome data (reported here) with that of Mo et al. (2016) revealed dissimilarities in photoreceptor gene expression profiles, which can be attributed to different methods of cell purification and genetic background of mouse lines. Further integration of transcriptome datasets with chromatin state of developing rods that we are generating and with the reported epigenomic data (Mo et al., 2016) would permit inference of functional interactions among the regulatory components in the “omics” landscape of photoreceptors and help in formulating testable hypotheses. Our studies also offer a framework for identifying convergent networks (Menche et al., 2015) for photoreceptor degeneration. The identification of a manageable number of common networks, upon which converge independent pathways affected by gene mutations, would facilitate the discovery of drug targets and therapies (Swaroop et al., 2010).

EXPERIMENTAL PROCEDURES

Isolation of Mouse Rod Photoreceptors

Animal Care and Use Committee of the NEI approved all procedures that involved mice. GFP-positive cells were purified from the retina of *Nrlp*-GFP and *Nrlp*-GFP;Nrl-KO mice (Akimoto et al., 2006). The retinas were dissected in Hank’s Balanced Salt Solution (Life Technologies, NY, USA) at different postnatal stages of development, followed by dissociation in Accutase (Life Technologies) at 37°C for 10 min with constant agitation. After removing cell clumps by filtration, 4 ml of PBS (Life Technologies) was added to dissociated cells. The cells were collected by centrifugation at 800*g* for 5 min and re-suspended in 1 ml of PBS. GFP-positive cells were isolated by fluorescence-activated cell sorting (FACS) using FACS Aria II (Becton Dickinson, CA, USA) with a stringent precision setting to maximize the purity of the sorted cells. The purity of isolated GFP-positive cells

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was evaluated by resorting. After sorting, cells were processed with TRIzol LS (Invitrogen) following manufacturer's instruction. For exon microarrays, *Nrl*p-GFP mouse retinas were treated in a similar manner, except that cells were dissociated by incubation in papain (Worthington Biochemical, NJ, USA) supplemented with DNase I, superoxide dismutase, catalase, D-α-tocopherol acetate, and gentamycin at 28°0 for 8 mi n. GFP-positive cells were collected into RNAProtect (Qiagen, Hilden, Germany).

Expression Data Generation and Analysis

Microarray data was generated using the WT-Ovation RNA Amplification System (NuGEN Technologies, San Carlos, CA) and hybridized to GeneChip® Mouse Exon 1.0 ST Arrays (Affymetrix). Gene-level core RMA intensity values for each array were collected by Expression Console Software (Affymetrix) for mm8 annotation. Strand-specific RNA-seq data was generated using TruSeq RNA Sample Prep Kit-v2 (Illumina), as described (Brooks et al., 2012), and 76 base single-end reads were generated on Genome Analyzer IIx platform (Illumina). Transcript levels were quantified using Ensembl v73 transcriptome annotation, as described (Kaewkhaw et al., 2015).

Details of microarray hybridization, strand-specific RNA-seq, differential gene expression, RRBS DNA methylation, and bioinformatic analysis are described in the Supplemental Material.

In Situ Hybridization (ISH)

C57BL/6 wild type retina at PN6 and PN28 was fixed with 4% paraformaldehyde in PBS at 4°C overnight and transferred to 30% sucrose for overnight incubation at 4°C. The retina was then embedded in OCT mounting medium, cryosectioned at 12 μm thickness and mounted on SuperFrost Plus glass slides. Retinal sections were then processed for florescent ISH using RNAscope technology (Advanced Cell Diagnostics, Hayward, CA), following the manufacturer's instructions with minor modifications. Probes were designed to recognize approximately 1000 nucleotides within each transcript (Supplemental Table S7). Briefly, all probes were assayed using the RNAscope Multiplex Fluorescent kit. Slides were washed in PBS and incubated in Pretreat 2 solution at 1000°C for 3 min, followed by incubation in Pretreat 3 solution at 40°C for 30 mi n. RNAscope probes were then hybridized at 40°C for 2 hours, followed by amplification step s according to the manufacturer's instructions. The signal was visualized and captured using a Zeiss 780 confocal microscope (Zeiss, Germany).

In Vivo Electroporation of shRNA

Custom shRNAs were generated against the 3'UTR of the novel TCONS_00068375 gene according to the Addgene TRC cloning vector protocol (https://www.addgene.org/tools/protocols/plko/#C). Briefly, oligonucleotides containing the shRNA insert were annealed together, and the annealed oligonucleotides and the cloning vector (Addgene plasmid 10878) were digested with AgeI-HF and EcoRI-HF. Digested oligonucleotides and gel-purified, digested plasmid were then ligated together. Constructed custom shRNA plasmids were transformed, amplified, and sequenced prior to *in vivo* electroporation. *In vivo* electroporation of the shRNA constructs were performed as described previously (Matsuda and Cepko 2004). One eye of P0 CD1 mice was injected subretinally with 0.2 μl of solution

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

containing shRNA or control constructs and reporter plasmids, followed by application of 5 times of 80V pulses for 50 msec with 950 msec interval to deliver DNA into the outer portion of the retina. To track total transfected cells as well as transfected rods, 500 ng/ul of each fluorescent reporter, *CAG*p-mCherry and *Nrl*p-GFP, respectively, was mixed with 2000 ng/μl of the control (Addgene plasmid 10879) or custom shRNA constructs. Two distinct shRNA constructs targeting the same gene and the control vector were injected in a single litter. Electroporated mice were sacrificed at P21, and the harvested retinas were fixed for 1 hr in 4% paraformaldehyde, cryopreserved in 30% sucrose, sectioned at 12 μm, and stained with DAPI. Fluorescence Images were taken at 20x magnification on the Zeiss 700 confocal microscope.

Statistical Analysis

Statistical analysis of variances between two different experimental groups was conducted with Tukey's post hoc comparison test using SPSS (SPSS, Chicago, IL, USA). All experiments were repeated at least three times. The levels were considered significant at $p < 0.05$ (*) and very significant at $p < 0.01$ (**), or not significant (n.s.).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Julie Laux of NEI flow-cytometry core for assistance. This research was supported by Intramural Research Program of the NEI (EY000450 and EY000456) and the National Research Foundation of Korea (2015R1C1A1A01055466 and 2016R1A4A1008035) and utilized the high-performance computational capabilities of the Biowulf Linux cluster at NIH (http://biowulf.nih.gov).

REFERENCES

Adler R, Hatlee M. Plasticity and differentiation of embryonic retinal cells after terminal mitosis. Science. 1989; 243:391–393. [PubMed: 2911751]

Akimoto M, Cheng H, Zhu D, Brzezinski JA, Khanna R, Filippova E, Oh EC, Jing Y, Linares JL, Brooks M, et al. Targeting of GFP to newborn rods by Nrl promoter and temporal expression profiling of flow-sorted photoreceptors. Proc Natl Acad Sci U S A. 2006; 103:3890–3895. [PubMed: 16505381]

Allan DW, Thor S. Transcriptional selectors, masters, and combinatorial codes: regulatory principles of neural subtype specification. Wiley Interdiscip Rev Dev Biol. 2015; 4:505–528. [PubMed: 25855098]

Blanks JC, Adinolfi AM, Lolley RN. Synaptogenesis in the photoreceptor terminal of the mouse retina. J Comp Neurol. 1974; 156:81–93. [PubMed: 4836656]

Brooks MJ, Rajasimha HK, Swaroop A. Retinal transcriptome profiling by directional next-generation sequencing using 100 ng of total RNA. Methods Mol Biol. 2012; 884:319–334. [PubMed: 22688717]

Brzezinski, JAt; Uoon Park, K.; Reh, TA. Blimp1 (Prdm1) prevents re-specification of photoreceptors into retinal bipolar cells by restricting competence. Dev Biol. 2013; 384:194–204. [PubMed: 24125957]

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013; 10:1213–1218. [PubMed: 24097267]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet. 2002; 3:285–298. [PubMed: 11967553]

Carter-Dawson LD, LaVail MM. Rods and cones in the mouse retina. II. Autoradiographic analysis of cell generation using tritiated thymidine. J Comp Neurol. 1979; 188:263–272. [PubMed: 500859]

Chakravarti A, Clark AG, Mootha VK. Distilling pathophysiology from complex disease genetics. Cell. 2013; 155:21–26. [PubMed: 24074858]

Chen S, Wang QL, Nie Z, Sun H, Lennon G, Copeland NG, Gilbert DJ, Jenkins NA, Zack DJ. Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. Neuron. 1997; 19:1017–1030. [PubMed: 9390516]

Cheng H, Aleman TS, Cideciyan AV, Khanna R, Jacobson SG, Swaroop A. In vivo function of the orphan nuclear receptor NR2E3 in establishing photoreceptor identity during mammalian retinal development. Hum Mol Genet. 2006; 15:2588–2602. [PubMed: 16868010]

Consortium EP. The ENCODE (ENCyclopedia Of DNA Elements) Project. Science. 2004; 306:636–640. [PubMed: 15499007]

Corbo JC, Lawrence KA, Karlstetter M, Myers CA, Abdelaziz M, Dirkes W, Weigelt K, Seifert M, Benes V, Fritsche LG, et al. CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. Genome Res. 2010; 20:1512–1525. [PubMed: 20693478]

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, et al. The Pfam protein families database. Nucleic Acids Res. 2010; 38:D211–D222. [PubMed: 19920124]

Fritsche LG, Igl W, Bailey JN, Grassmann F, Sengupta S, Bragg-Gresham JL, Burdon KP, Hebbring SJ, Wen C, Gorski M, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet. 2016; 48:134–143. [PubMed: 26691988]

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. Nature. 2015; 526:68–74. [PubMed: 26432245]

Green CB, Besharse JC. Retinal circadian clocks and control of retinal physiology. J Biol Rhythms. 2004; 19:91–102. [PubMed: 15038849]

Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc. 2011; 6:468–481. [PubMed: 21412275]

Guillemot F. Spatial and temporal specification of neural fates by transcription factor codes. Development. 2007; 134:3771–3780. [PubMed: 17898002]

Hao H, Kim DS, Klocke B, Johnson KR, Cui K, Gotoh N, Zang C, Gregorski J, Gieser L, Peng W, et al. Transcriptional regulation of rod photoreceptor homeostasis revealed by in vivo NRL targetome analysis. PLoS Genet. 2012; 8:e1002649. [PubMed: 22511886]

Hao H, Tummala P, Guzman E, Mali RS, Gregorski J, Swaroop A, Mitton KP. The transcription factor neural retina leucine zipper (NRL) controls photoreceptor-specific expression of myocyte enhancer factor Mef2c from an alternative promoter. J Biol Chem. 2011; 286:34893–34902. [PubMed: 21849497]

Hao H, Veleri S, Sun B, Kim DS, Keeley PW, Kim JW, Yang HJ, Yadav SP, Manjunath SH, Sood R, et al. Regulation of a novel isoform of Receptor Expression Enhancing Protein REEP6 in rod photoreceptors by bZIP transcription factor NRL. Hum Mol Genet. 2014; 23:4260–4271. [PubMed: 24691551]

Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, van de Lagemaat LN, Smith KA, Ebbert A, Riley ZL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012; 489:391–399. [PubMed: 22996553]

Jia L, Oh EC, Ng L, Srinivas M, Brooks M, Swaroop A, Forrest D. Retinoid-related orphan nuclear receptor RORbeta is an early-acting factor in rod photoreceptor development. Proc Natl Acad Sci U S A. 2009; 106:17534–17539. [PubMed: 19805139]

Kaewkhaw R, Kaya KD, Brooks M, Homma K, Zou J, Chaitankar V, Rao M, Swaroop A. Transcriptome Dynamics of Developing Photoreceptors in Three-Dimensional Retina Cultures

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Recapitulates Temporal Sequence of Human Cone and Rod Differentiation Revealing Cell Surface Markers and Gene Networks. Stem Cells. 2015; 33:3504–3518. [PubMed: 26235913]

Katoh K, Omori Y, Onishi A, Sato S, Kondo M, Furukawa T. Blimp1 suppresses Chx10 expression in differentiating retinal photoreceptor precursors to ensure proper photoreceptor development. J Neurosci. 2010; 30:6515–6526. [PubMed: 20463215]

Kim JW, Yang HJ, Oel AP, Brooks MJ, Jia L, Plachetzki DC, Li W, Allison WT, Swaroop A. Recruitment of Rod Photoreceptors from Short Wavelength Sensitive Cones during the Evolution of Nocturnal Vision in Mammals. Dev Cell. 2016 in press.

Kooragayala K, Gotoh N, Cogliati T, Nellissery J, Kaden TR, French S, Balaban R, Li W, Covian R, Swaroop A. Quantification of Oxygen Consumption in Retina Ex Vivo Demonstrates Limited Reserve Capacity of Photoreceptor Mitochondria. Invest Ophthalmol Vis Sci. 2015; 56:8428–8436. [PubMed: 26747773]

Lamb TD, Collin SP, Pugh EN Jr. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. Nat Rev Neurosci. 2007; 8:960–976. [PubMed: 18026166]

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. [PubMed: 19114008]

Li B, Fillmore N, Bai Y, Collins M, Thomson JA, Stewart R, Dewey CN. Evaluation of de novo transcriptome assemblies from RNA-Seq data. Genome Biol. 2014; 15:553. [PubMed: 25608678]

Liu MM, Zack DJ. Alternative splicing and retinal degeneration. Clin Genet. 2013; 84:142–149. [PubMed: 23647439]

Livne-Bar I, Pacal M, Cheung MC, Hankin M, Trogadis J, Chen D, Dorval KM, Bremner R. Chx10 is required to block photoreceptor differentiation but is dispensable for progenitor proliferation in the postnatal retina. Proc Natl Acad Sci U S A. 2006; 103:4988–4993. [PubMed: 16547132]

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]

Mattson MP, Magnus T. Ageing and neuronal vulnerability. Nat Rev Neurosci. 2006; 7:278–294. [PubMed: 16552414]

McMahon DG, Iuvone PM, Tosini G. Circadian organization of the mammalian retina: from gene regulation to physiology and diseases. Prog Retin Eye Res. 2014; 39:58–76. [PubMed: 24333669]

Mears AJ, Kondo M, Swain PK, Takada Y, Bush RA, Saunders TL, Sieving PA, Swaroop A. Nrl is required for rod photoreceptor development. Nat Genet. 2001; 29:447–452. [PubMed: 11694879]

Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, Barabasi AL. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. Science. 2015; 347:1257601. [PubMed: 25700523]

Mitton KP, Swain PK, Chen S, Xu S, Zack DJ, Swaroop A. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. J Biol Chem. 2000; 275:29794–29799. [PubMed: 10887186]

Mo A, Luo C, Davis FP, Mukamel EA, Henry GL, Nery JR, Urich MA, Picard S, Lister R, Eddy SR, et al. Epigenomic landscapes of retinal rods and cones. Elife. 2016; 5:e11613. [PubMed: 26949250]

Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, Albu M, Weirauch MT, Radovani E, Kim PM, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat Biotechnol. 2015; 33:555–562. [PubMed: 25690854]

Ng L, Lu A, Swaroop A, Sharlin DS, Swaroop A, Forrest D. Two transcription factors can direct three photoreceptor outcomes from rod precursor cells in mouse retinal development. J Neurosci. 2011; 31:11118–11125. [PubMed: 21813673]

Nishida A, Furukawa A, Koike C, Tano Y, Aizawa S, Matsuo I, Furukawa T. Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. Nat Neurosci. 2003; 6:1255–1263. [PubMed: 14625556]

Obata S, Usukura J. Morphogenesis of the photoreceptor outer segment during postnatal development in the mouse (BALB/c) retina. Cell Tissue Res. 1992; 269:39–48. [PubMed: 1423483]

Oh EC, Cheng H, Hao H, Jia L, Khan NW, Swaroop A. Rod differentiation factor NRL activates the expression of nuclear receptor NR2E3 to suppress the development of cone photoreceptors. Brain Res. 2008; 1236:16–29. [PubMed: 18294621]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Oh EC, Khan N, Novelli E, Khanna H, Strettoi E, Swaroop A. Transformation of cone precursors to functional rod photoreceptors by bZIP transcription factor NRL. Proc Natl Acad Sci U S A. 2007; 104:1679–1684. [PubMed: 17242361]

Onishi A, Peng GH, Poth EM, Lee DA, Chen J, Alexis U, de Melo J, Chen S, Blackshaw S. The orphan nuclear hormone receptor ERRbeta controls rod photoreceptor survival. Proc Natl Acad Sci USA. 2010; 107:11579–11584. [PubMed: 20534447]

Peter IS, Davidson EH. Evolution of gene regulatory networks controlling body plan development. Cell. 2011; 144:970–985. [PubMed: 21414487]

Roger JE, Hiriyanna A, Gotoh N, Hao H, Cheng DF, Ratnapriya R, Kautzmann MA, Chang B, Swaroop A. OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. J Clin Invest. 2014; 124:631–643. [PubMed: 24382353]

Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. 2004; 32:D91–D94. [PubMed: 14681366]

Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet. 2013; 14:618–630. [PubMed: 23897237]

Siegert S, Cabuy E, Scherf BG, Kohler H, Panda S, Le YZ, Fehling HJ, Gaidatzis D, Stadler MB, Roska B. Transcriptional code and disease map for adult retinal cell types. Nat Neurosci. 2012; 15:487–495. S481-482. [PubMed: 22267162]

Skokowa J, Klimiankou M, Klimenkova O, Lan D, Gupta K, Hussein K, Carrizosa E, Kusnetsova I, Li Z, Sustmann C, et al. Interactions among HCLS1, HAX1 and LEF-1 proteins are essential for G-CSF-triggered granulopoiesis. Nat Med. 2012; 18:1550–1559. [PubMed: 23001182]

Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, Andrews SR, Stegle O, Reik W, Kelsey G. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods. 2014; 11:817–820. [PubMed: 25042786]

Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012; 13:613–626. [PubMed: 22868264]

Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008; 9:465–476. [PubMed: 18463664]

Swaroop A, Kim D, Forrest D. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. Nat Rev Neurosci. 2010; 11:563–576. [PubMed: 20648062]

Tebbenkamp AT, Willsey AJ, State MW, Sestan N. The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. Curr Opin Neurol. 2014; 27:149–156. [PubMed: 24565942]

Telley L, Govindan S, Prados J, Stevant I, Nef S, Dermitzakis E, Dayer A, Jabaudon D. Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. Science. 2016; 351:1443–1446. [PubMed: 26940868]

Thakurela S, Sahu SK, Garding A, Tiwari VK. Dynamics and function of distal regulatory elements during neurogenesis and neuroplasticity. Genome Res. 2015; 25:1309–1324. [PubMed: 26170447]

Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, Amit I, Gnirke A, Meissner A. Transcription factor binding dynamics during human ES cell differentiation. Nature. 2015; 518:344–349. [PubMed: 25693565]

van de Leemput J, Boles NC, Kiehl TR, Corneo B, Lederman P, Menon V, Lee C, Martinez RA, Levi BP, Thompson CL, et al. CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. Neuron. 2014; 83:51–68. [PubMed: 24991954]

Wan J, Masuda T, Hackler L Jr, Torres KM, Merbs SL, Zack DJ, Qian J. Dynamic usage of alternative splicing exons during mouse retina development. Nucl Acids Res. 2011; 39:7920–7930. [PubMed: 21724604]

Wang S, Sengel C, Emerson MM, Cepko CL. A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. Dev Cell. 2014; 30:513–527. [PubMed: 25155555]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Wong JJ, Ritchie W, Ebner OA, Selbach M, Wong JW, Huang Y, Gao D, Pinello N, Gonzalez M, Baidya K, et al. Orchestrated intron retention regulates normal granulocyte differentiation. Cell. 2013; 154:583–595. [PubMed: 23911323]

Wright AF, Chakarova CF, Abd El-Aziz MM, Bhattacharya SS. Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. Nat Rev Genet. 2010; 11:273–284. [PubMed: 20212494]

Yadav SP, Hao H, Yang HJ, Kautzmann MA, Brooks M, Nellissery J, Klocke B, Seifert M, Swaroop A. The transcription-splicing protein NonO/p54nrb and three NonO-interacting proteins bind to distal enhancer region and augment rhodopsin expression. Hum Mol Genet. 2014; 23:2132–2144. [PubMed: 24301678]

Yang HJ, Ratnapriya R, Cogliati T, Kim JW, Swaroop A. Vision from next generation sequencing: multi-dimensional genome-wide analysis for producing gene regulatory networks underlying retinal development, aging and disease. Prog Retin Eye Res. 2015; 46:1–30. [PubMed: 25668385]

Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV. Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. 2012; 26:1209–1223. [PubMed: 22661231]

Yoshida S, Mears AJ, Friedman JS, Carter T, He S, Oh E, Jing Y, Farjo R, Fleury G, Barlow C, et al. Expression profiling of the developing and mature Nrl−/− mouse retina: identification of retinal disease candidates and transcriptional regulatory targets of Nrl. Hum Mol Genet. 2004; 13:1487–1503. [PubMed: 15163632]

Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci. 2014; 34:11929–11947. [PubMed: 25186741]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

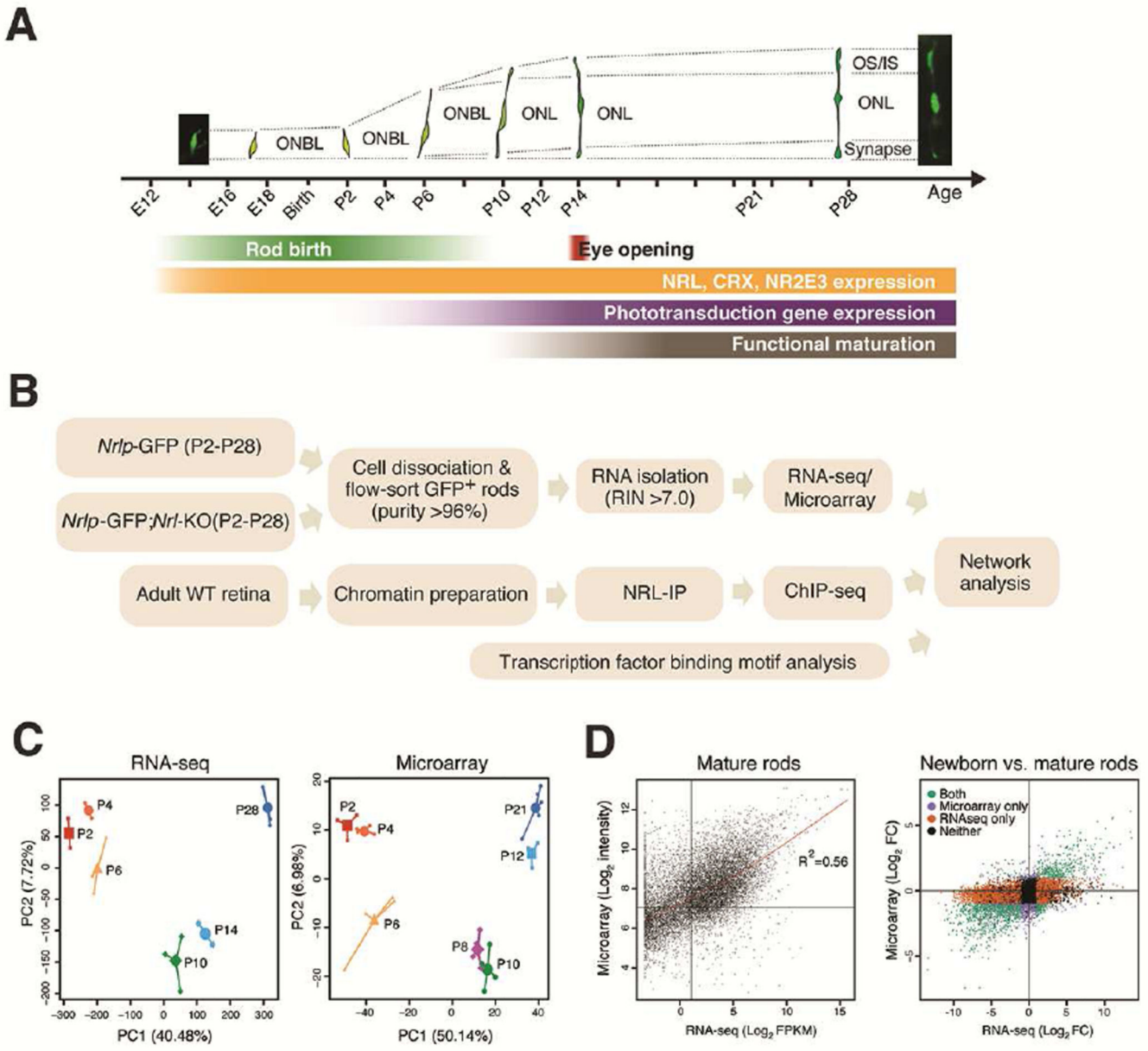


Figure 1. Study Design and Data Generation

(**A**) Time course of rod photoreceptor differentiation in mouse retina. Morphologies of developing and mature rods are shown at indicated developmental stages, with major events summarized below. ONBL, outer neuroblastic layer; ONL, outer nuclear layer; OS/IS, outer segment/inner segment. (**B**) Flowchart of integrated transcriptome and network analysis. WT, wild type; RIN, RNA integration number; IP, immunoprecipitation. (**C, D**) Comparison of rod transcriptomes generated by directional RNA-seq and exon microarray. (**C**) Principal component analysis (PCA) of directional RNA-seq (left panel) and microarray (right panel) data. The percentages indicate the amount of variation attributed to each principal component. Small shapes indicate individual samples, and larger shapes show the centroid

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of each grouping. (**D**) Correlation between RNA-seq and microarray data. RNA-seq and microarray data were de-duplicated and merged based on gene symbols. The highest FC representative from the RNA-seq transcriptome data was used as the gene representative for RNA-seq. Left Panel: Expression values (in $\log_2$ scale) of individual genes of mature rods (i.e., P28 in RNA-seq and P21 in microarray) were compared with each other using a scatter plot. Log_2 FPKM of one and $\log_2$ signal intensity of seven was used as a cut-off of actively transcribed genes in RNA-seq and microarray, respectively. The lowess regression line is shown in red, and the Pearson's correlation coefficient is indicated. Right Panel: Differentially expressed genes between newborn (P2) and mature rods (P21 for microarray and P28 for RNA-seq) were identified as having >2 fold change and FDR<0.05 in microarray data or <0.01 in RNA-seq data. The scatter plot shows genes that are significantly differentially expressed genes in both analyses, microarray only, and RNA-seq only (green, purple, and orange dots, respectively). Black dots indicate genes with no significant change in expression in either analysis. FPKM, fragments per kilobase of exon model per millions of reads; FC, fold change; FDR, Benjamini-Hochberg false discovery rate.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

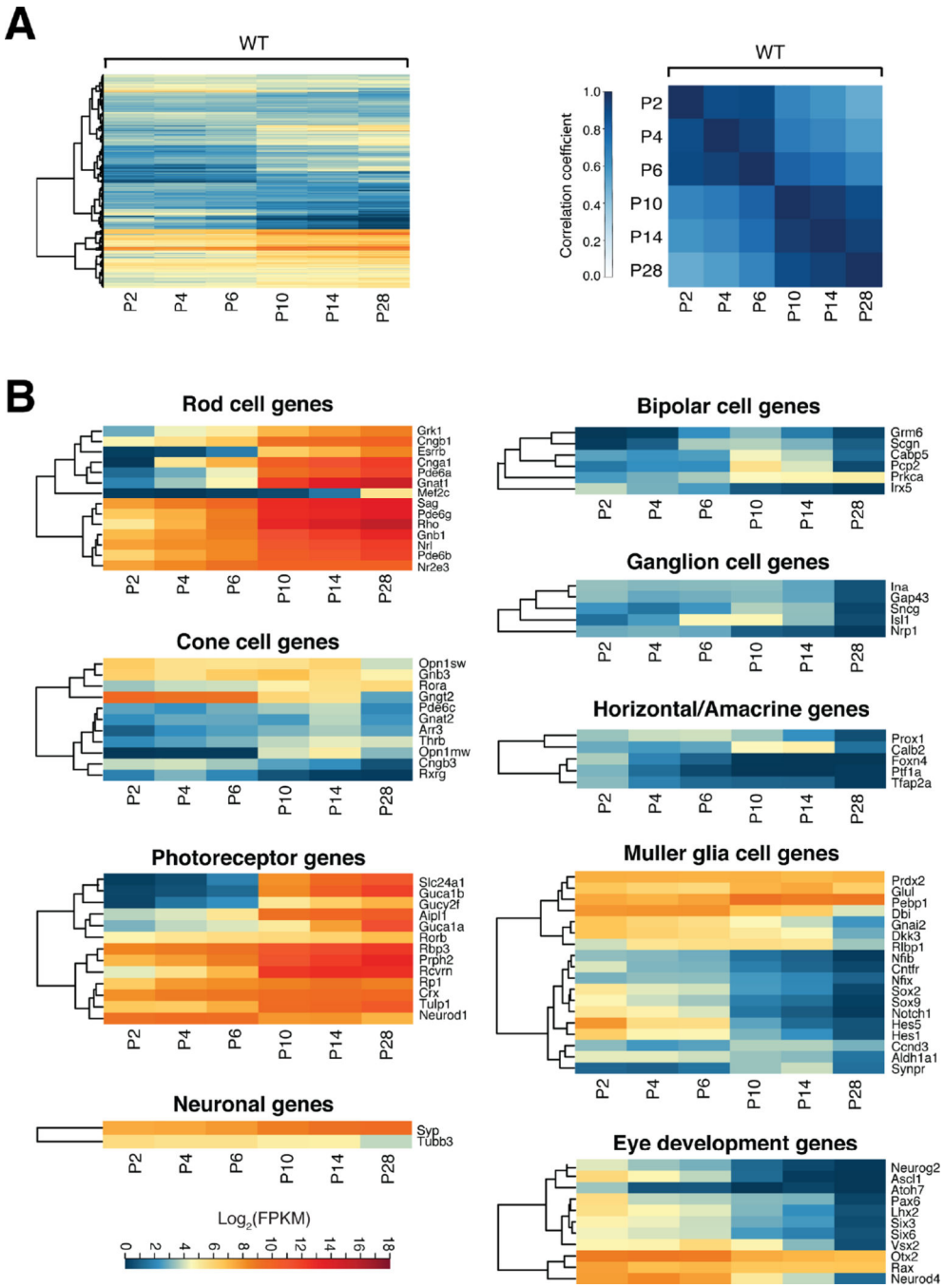


Figure 2. Transcriptome Dynamics during Rod Differentiation

(**A**) Hierarchically-clustered heat map of transcripts (a total of 11,146 transcripts that were expressed ≥5 FPKM at any time point in WT) during rod photoreceptor development (left panel) and correlation coefficient between every possible pair of transcriptomes at various differentiation stages calculated from directional RNA-seq data and presented in a matrix (right panel). A sharp change in transcriptome landscape is apparent between P6 and P10. (**B**) Dynamic expression patterns of genes specific for photoreceptors (14 rod genes, 11 cone genes and 13 genes expressed in both) and other retinal cell types during rod differentiation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In heat maps, the average expression at each time point is plotted in $\log_2$ scale, and only those transcripts that were expressed at ≥5 in all replicates of at least one time point were included. Color scale is indicated in the bottom.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

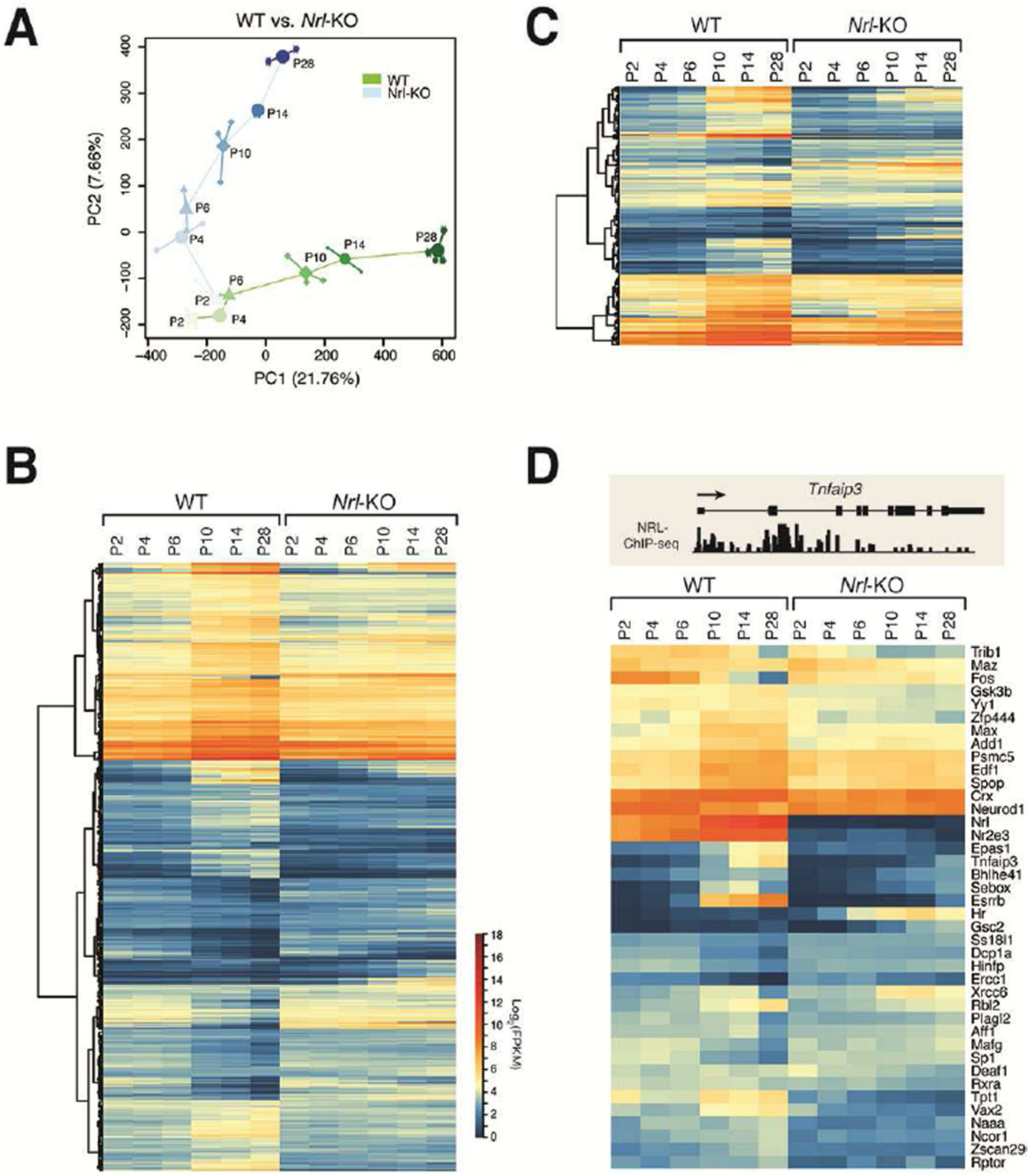


Figure 3. Transcriptome regulation by NRL

(**A**) PCA analysis of time series RNA-seq data from *Nrl*-GFP+ wild type rod photoreceptors and S-cone like photoreceptors from the *Nrl*-KO retina. A significant change in transcriptome landscape was detected in GFP+ *Nrl*-KO photoreceptors. (**B**) Heat map of significantly, differentially expressed protein-coding transcripts. Color scale is indicated. (**C**) Heat map of differentially expressed genes containing NRL ChIP-seq peaks. An integrated analysis was performed combining differential gene expression data with the NRL ChIP-seq (Hao et al., 2012). The DE transcripts having at least one NRL ChIP-seq peak were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

considered putative, direct transcriptional targets and were included in the heat map. (**D**) Identification of transcription regulatory proteins that are putative NRL direct targets. Shown on the top is a genomic locus of one of the differentially expressed genes, *Tnfaip3*, and NRL ChIP-seq coverage plot.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

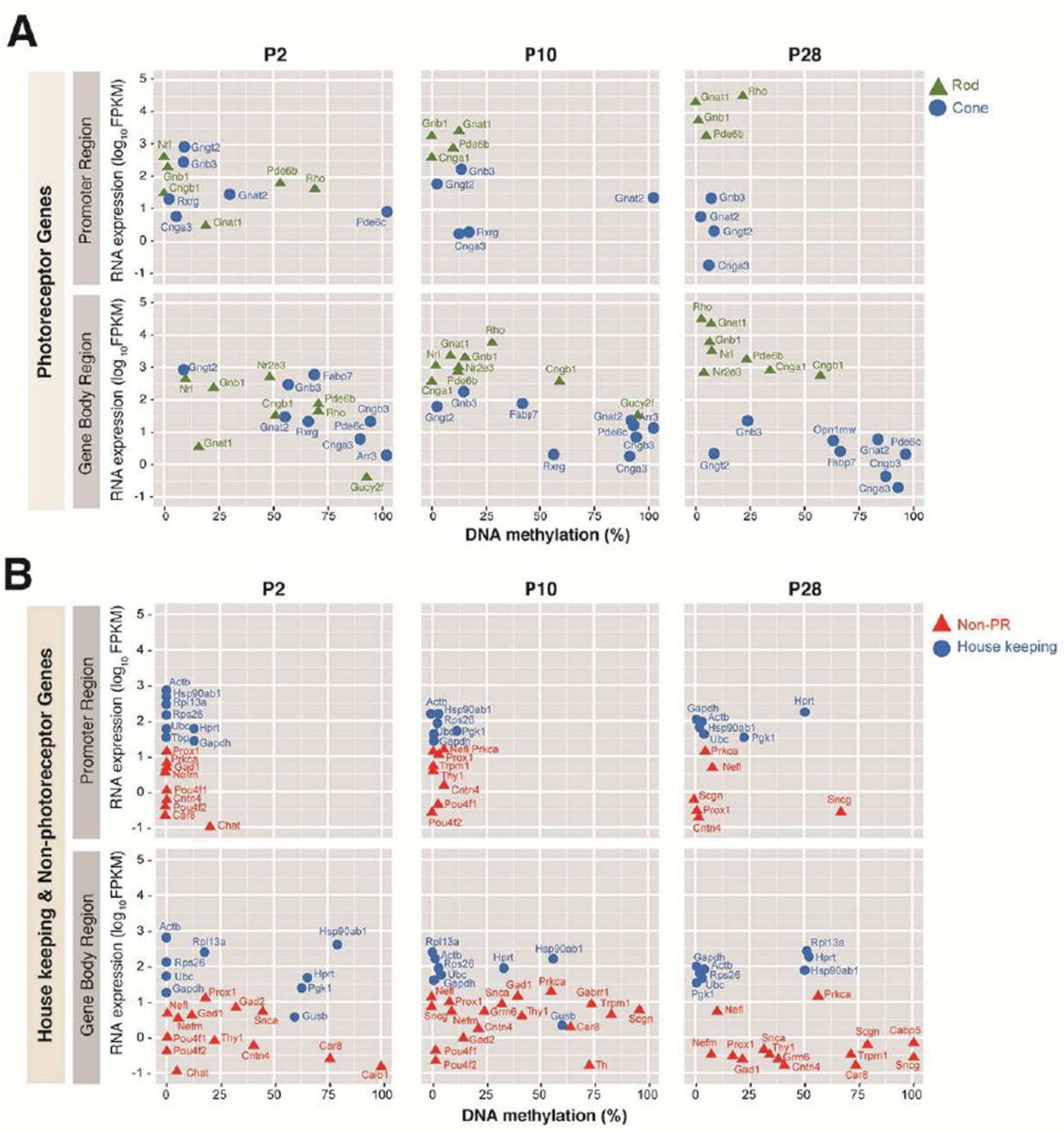


Figure 4. DNA Methylation of Photoreceptor Genes

(**A**) Correlation between DNA methylation level and gene expression of rod- and cone-specific genes. Degree of DNA methylation (%) in promoter or gene body region of individual rod- and cone-specific genes (green triangle and blue circle, respectively) are plotted against gene expression level (in $\log_{10}$FPKM). Please note that some promoters and gene bodies were not covered in RRBS analysis and thus not included in the plot. (**B**) Plotted is correlation between DNA methylation level (%) in promoter or in gene body

region and gene expression (in $\log_{10}$FPKM) of non-photoreceptor retinal genes (red triangle) and housekeeping genes (blue circle).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

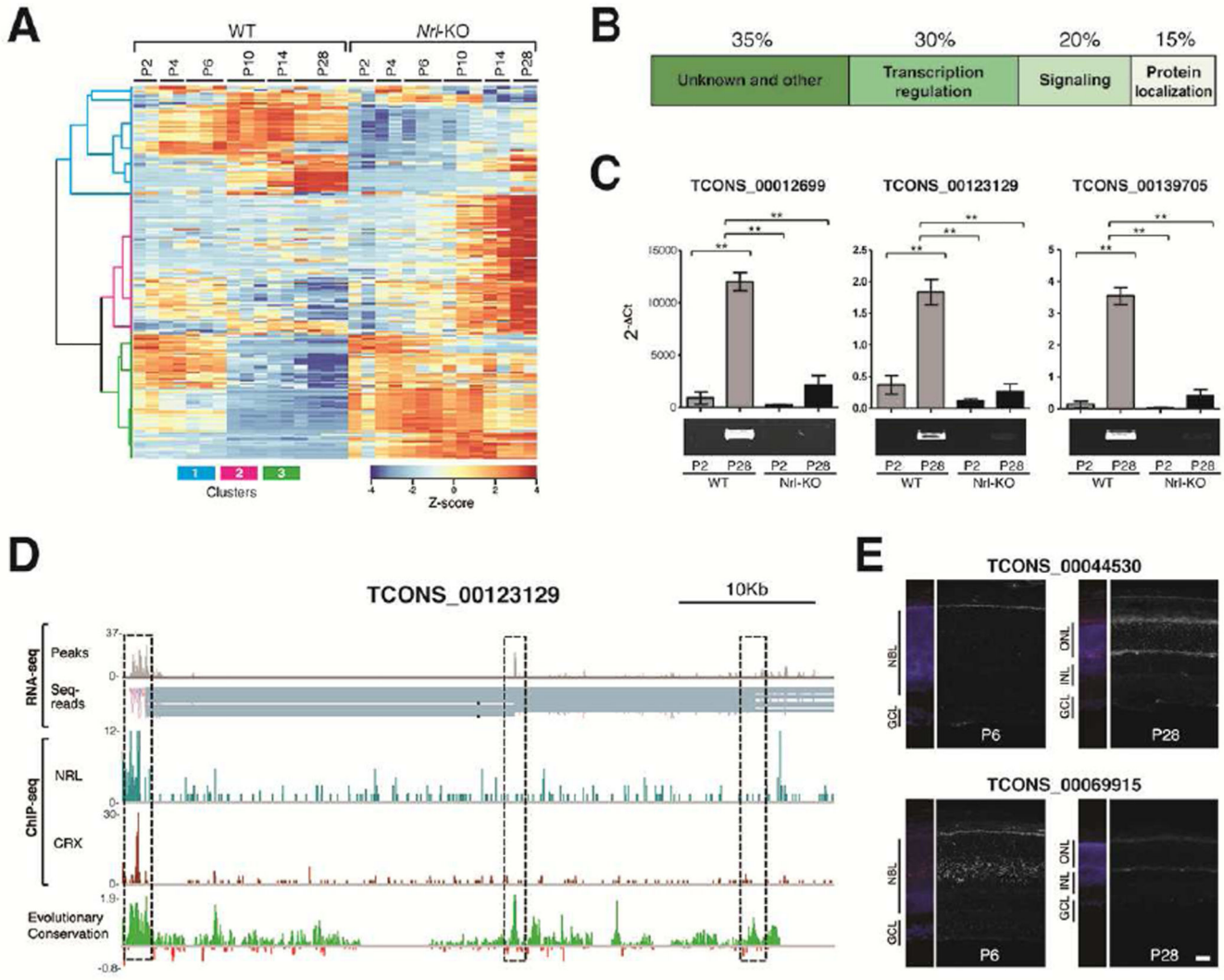


Figure 5. ***In-silico De Novo*** **Analysis and Validation of Putative Protein Coding Transcripts**

(**A**) Heat map of standardized $\log_2$ FPKM values (with one offset) of 222 previously un-annotated protein-coding transcripts (NPCs) across all time points in rods from wild type and S-cone-like photoreceptors from *Nrl*-KO retina. Individual replicates are shown to display the degree of variability intrinsic to this type of analysis. The three largest clusters in the dendrogram are highlighted in blue (1), magenta (2), and green (3), respectively. (**B**) Functional stratification of conserved protein domain identified by HMMER. Identified domains were grouped according to their described function in Pfam database. The relative frequency of occurrence was calculated for domains observed at least 10 times among the significant protein domain hits. (**C**) qRT-PCR validation of three select NPC transcripts. Relative enrichment of each transcript compared with expression of housekeeping gene *Actb* was plotted for P2 and P28 wild type and *Nrl*-KO retinas. Double asterisks indicate statistically significant changes. Agarose gel images are shown below. (**D**) Schematic representation of a NPC, TCONS_00123129. NRL and CRX ChIP-seq peaks are located in close proximity to the TSS of this transcript and placental mammalian evolutionary

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

conservation peaks (as available from USCS genome browser) overlap with the predicted gene structure, highlighted by a dashed box. Histograms for RNA-seq and NRL- and CRX-ChlP-seq indicate the coverage of all aligned reads across the genomic area. Below the RNA-seq histogram are shown the individual sequence reads. Red or blue blocks indicate discrete reads in antisense or sense orientation, respectively, with thin blue lines marking the splice portions of the reads. **(E)** *In-situ* hybridization of the previously un-annotated transcripts TCONS_00044530 and TCONS_00069915 at P6 and P28 in wild type retina. The signal for both transcripts follows the dynamic expression pattern seen in panel A and Figure S3. Blue, DAPI nuclear stain; Red, *in situ* signal in merged images; White, *in situ* signal in single channel images. Scale bar, 50 μm. GCL, ganglion cell layer; INL, inner nuclear layer; NBL, neuroblastic layer; ONL, outer nuclear layer.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

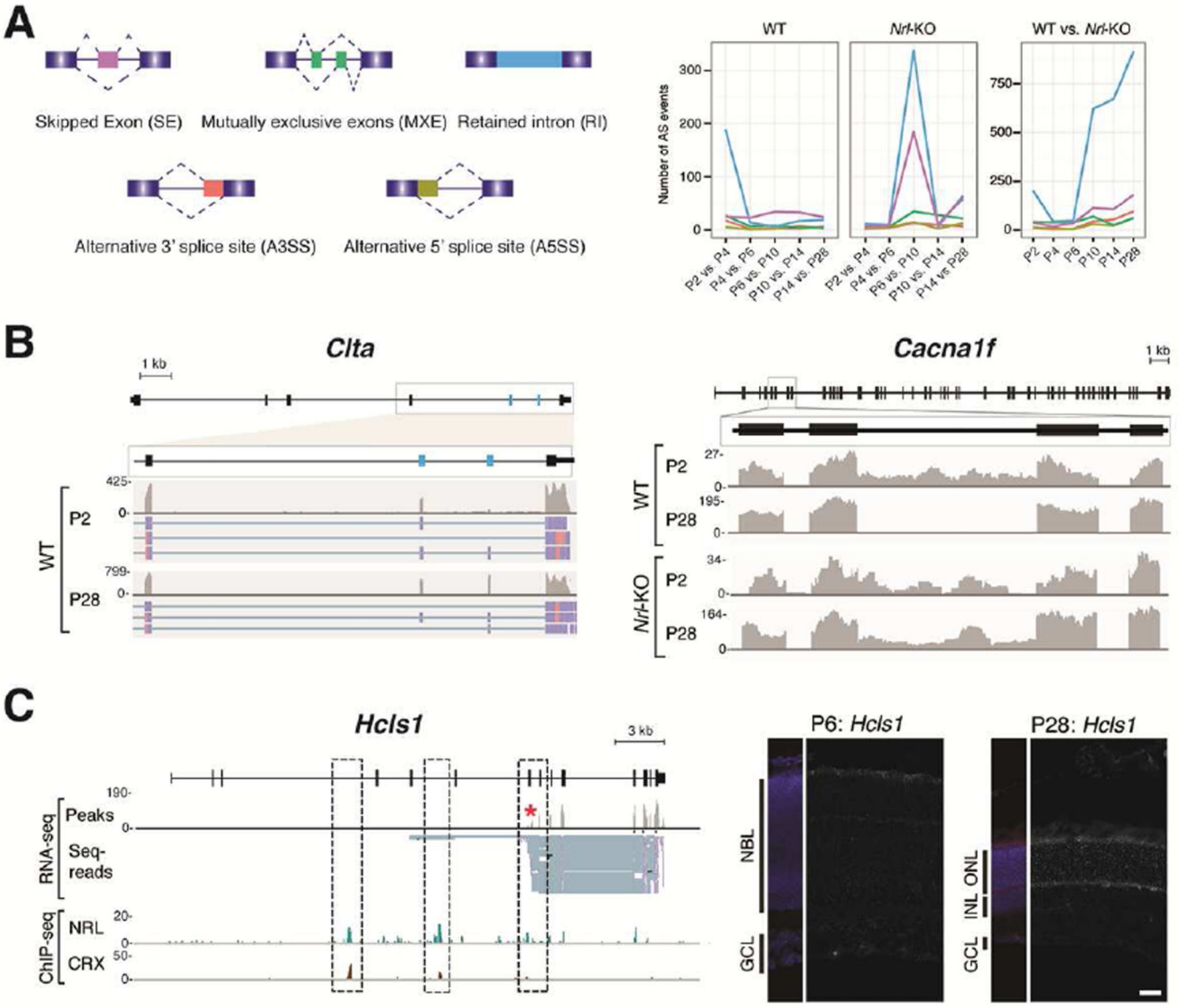


Figure 6. Alternative Splicing and Promoter Usage Events Identified during Rod Differentiation

(A) Types of alternative splicing events detected and their frequencies in time-wise and group-wise comparisons. Splicing events are color coded to match the relevant graphs. **(B)** Differential splicing events during maturation of rods and *Nrl*-KO S-cone like photoreceptors. Shown is the genomic area of *Clta* gene (upper lane in the left panel), which comprises seven exons and six introns, and RNA-seq coverage plots of P2 and P28 wild type rods in the area of exons four through seven (bottom lanes in the left panel). Solid boxes and lines represent exons and introns, respectively, and untranslated regions are illustrated as thinner boxes. The fifth and sixth exons of *Clta* gene are differentially spliced and highlighted in blue. Individual spliced reads demonstrate the splice isoform that are generated from *Clta* at both time points. Translated exons, untranslated exonic regions and introns are indicated as thick boxes, thin boxes and lines, respectively. RNA-seq coverage plots are shown as gray histograms. For *Clta* gene, individual sequence reads (red and blue

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

boxes for antisense and sense orientation) with spliced area (thin blue lines) are also indicated. **(C)** Alternative promoter usage of *Hcls1*. RNA-seq reads show only partial coverage of known exons of *Hclsl* (exons 18–14), with another exon introduced upstream to exon 8 (red asterisk). Dashed boxes indicate NRL and CRX ChlP-seq peaks. *In situ* hybridization for a retinal isoform of *Hcls1* (probe designed to hybridize to approximately 1000 bases starting in the previously un-annotated exon) shows significant increase in expression from P6 to P28. DAPI nuclear stain is indicated in blue, and *in situ* signal is indicated in red and in white in merged or in single channel images, respectively. Scale bar, 50 μm. GCL, ganglion cell layer; INL, inner nuclear layer; NBL, neuroblastic layer; ONL, outer nuclear layer.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

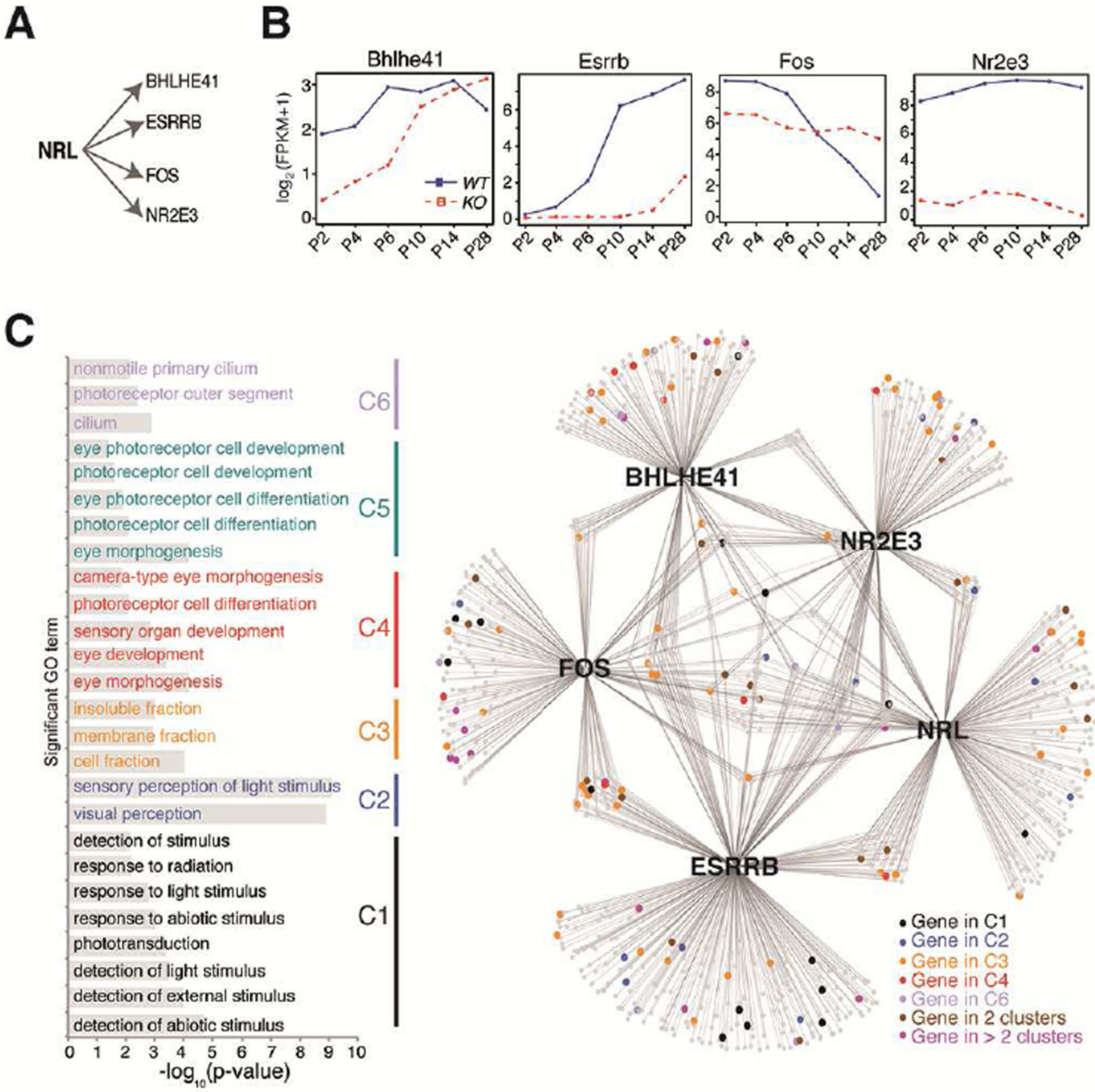


Figure 7. Gene Regulatory Network Analysis

(A) Regulatory hubs of NRL-centered GRN. **(B)** Temporal expression profile of secondary hubs. Expression values of four genes in rods (blue line) and *Nrl*-KO photoreceptors (red dashed line) are plotted in $\log_2$ scale for each time point. When multiple transcripts are expressed, expression level of the most highly expressed transcript is indicated. **(C)** NRL-centered network and highly represented GO terms among target genes of NRL and secondary hubs. Circles indicate individual nodes (i.e., regulators and target genes) with primary regulatory hub NRL and four other secondary hubs labeled, and lines indicate predicted interaction between regulatory hubs and targets. Some targets are regulated by one

regulatory gene, while others have regulatory interaction by two or more hubs. Target genes that belong to highly represented GO terms grouped into C1 through C5 were highlighted in indicated color.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript