# The Noncoding RNA *Taurine Upregulated Gene 1* Is Required for Differentiation of the Murine Retina

T.L. Young, T. Matsuda, and C.L. Cepko*
Howard Hughes Medical Institute
Department of Genetics
Harvard Medical School
77 Avenue Louis Pasteur
Boston, MA 02115

## Summary

**Background:** With the advent of genome-wide analyses, it is becoming evident that a large number of noncoding RNAs (ncRNAs) are expressed in vertebrates. However, of the thousands of ncRNAs identified, the functions of relatively few have been established.
**Results:** In a screen for genes upregulated by taurine in developing retinal cells, we identified a gene that appears to be a ncRNA. Taurine Upregulated Gene 1 (*TUG1*) is a spliced, polyadenylated RNA that does not encode any open reading frame greater than 82 amino acids in its full-length, 6.7 kilobase (kb) RNA sequence. Analyses of Northern blots and in situ hybridization revealed that *TUG1* is expressed in the developing retina and brain, as well as in adult tissues. In the newborn retina, knockdown of *TUG1* with RNA interference (RNAi) resulted in malformed or nonexistent outer segments of transfected photoreceptors. Immunofluorescent staining and microarray analyses suggested that this loss of proper photoreceptor differentiation is a result of the disregulation of photoreceptor gene expression.
**Conclusions:** A function for a newly identified ncRNA, *TUG1*, has been established. *TUG1* is necessary for the proper formation of photoreceptors in the developing rodent retina.

## Introduction

Taurine is a cysteine derivative that has been shown to be necessary for proper neural development (reviewed in [1]). In addition, taurine deficiency in cats, rodents, and primates has been linked to both a failure of photoreceptors to mature correctly and to photoreceptor degeneration in the adult retina [2–4]. In the rodent retina, taurine is present at high levels during development, and the addition of exogenous taurine to developing retinal cells promotes rod photoreceptor production [5]. The rod induction is due to signals mediated by the binding of taurine to glycine-receptor α2 and GABA(A) receptors [6]. We performed a screen to identify the genes that are regulated after signaling through these receptors (T.L.Y. et al., unpublished data). One of the most consistently and significantly upregulated genes is a novel gene, *Taurine Upregulated Gene 1* (*TUG1*).

Because taurine induces rod production during development, the genes that are upregulated by taurine are candidates for factors involved in photoreceptor development. Photoreceptors are highly specialized

*Correspondence: cepko@genetics.med.harvard.edu

cells that transduce visual information from light into a synaptic signal. Multiple proteins specific to the phototransduction cascade are localized to a membranous structure, the outer segment, where light is absorbed and much of the phototransduction cascade takes place. Transcription factors such as cone-rod homeobox (Crx) and neural retina leucine zipper (Nrl) have been found to be essential for the proper development of photoreceptors [7, 8]. Targeted deletion of Crx leads to a failure to form outer segments [7, 9]. As with Crx mutant animals, targeted deletion of Nrl leads to a loss of rod differentiation. However, in contrast to Crx null animals, there is an increase in the number of cone-like cells in the Nrl mutant [8]. In addition to *TUG1*, Nrl was one of the top genes that we identified in a microarray analysis as being upregulated by taurine (T.L.Y. et al., unpublished data). The current understanding of these gene regulatory networks is incomplete, and much remains to be learned about the generation and maintenance of these complex sensory cells.

*TUG1* does not appear to have significant homology to other genes in the mouse genome. However, a single highly conserved homolog of *TUG1* is present in human, rat, cow, and dog genomes. An interesting feature of *TUG1* RNA is that it does not have any open reading frames (ORFs) larger than 82 amino acids in any of the six reading frames of its full length, 6.7 kb of sequence. Therefore, *TUG1* either is a noncoding RNA (ncRNA) or encodes a very short peptide. Small peptides exist in vertebrates and are involved in a number of signaling processes. However, the majority of known small peptides, such as neuropeptides and peptide hormones, are processed from larger proteins (reviewed in [10]).

NcRNAs are emerging as a rapidly growing category of molecules involved in a number of regulatory processes. Traditionally, ncRNAs were thought to be limited to RNAs involved in the splicing and translation of messenger RNA (mRNA) into proteins. These included small RNA-polymerase-III-derived transcripts such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and U6 small nuclear RNAs (snRNA). Additionally, RNA polymerase II transcribes other small regulatory RNAs, including small nucleolar RNAs (snoRNAs) and four of the snRNAs. These classes of RNAs function in pre-RNA splicing, assembly of ribosomes, and RNA folding and cleavage. Micro RNAs (miRNAs), another class of ncRNAs, are 22 base pairs (bp) in length and are processed from larger RNA transcripts. MiRNAs have been shown to play roles in the regulation of cellular processes by specifically targeting certain mRNA species for degradation or translational repression.

A number of other RNA-polymerase-II transcripts have been identified that do not contain an ORF encoding greater than 100 amino acids and do not fit into one of the above categories. On the basis of analyses of coding-DNA (cDNA) libraries, it has been estimated that one-third to one-half of all transcripts in higher vertebrates do not contain a substantial ORF [11, 12]. In spite of this seeming abundance, only a small number of "large" RNA-polymerase-II ncRNAs have been char-

acterized experimentally (http://biobases.ibch.poznan.pl/ncRNA). One characteristic that should be noted is that many of the ncRNAs are defined as such on the basis of the fact that they do not contain a large ORF. It is difficult, however, to rule out the possibility that the small ORFs that are present encode functional proteins.

A limited number of these larger polymerase-II-transcribed ncRNAs have been shown to have a function in mammals. One of the few demonstrations of a ncRNA function in mammals is Xist, which is an essential component of X-chromosome inactivation. Xist is a large transcript that is transcribed by and physically coats the inactive X chromosome in both mice and humans (reviewed in [13]). After the molecular characterization of Xist, TsiX was identified as another ncRNA that regulates the function of Xist [14]. A third ncRNA that has been identified and proposed to have a function is H19. Loss of expression of the imprinted noncoding gene H19 has been linked to Wilms's tumors [15].

In this study, we describe the molecular characteristics of *TUG1*. *TUG1* is a gene that is upregulated by taurine and that is spliced and polyadenylated. In its full-length sequence, the largest ORF is only 82 amino acids, indicating that *TUG1* may be a ncRNA. In order to investigate the possible function of *TUG1* in photoreceptor development, we examined the expression of *TUG1* in the developing nervous system and used RNA interference (RNAi) technology to knock down *TUG1* in the developing retina.

## Results

### *TUG1* Genomic Organization and Predicted Full-Length cDNAs

*TUG1* was originally identified in a microarray screen for genes regulated in a heterogeneous culture of postnatal-day-0 (P0) retinal cells upon addition of taurine for 4 hr (T.L.Y. et al., unpublished data). This microarray consisted of approximately 12,000 brain and retinal cDNAs isolated by Dr. Bento Soares (http://genome.uiowa.edu) and 400 additional cDNAs representing genes of interest in our laboratory. *TUG1* upregulation was confirmed by quantitative reverse-transcription polymerase chain reaction (RT-PCR) to be approximately 3-fold. Because the observation that taurine stimulates rod production through activation of glycine and GABA receptors, which leads to depolarization of immature retinal cells, the regulation of *TUG1* by glycine, GABA, and KCl also was tested through microarray analyses. *TUG1* also was upregulated by these factors 4 hr after its addition to P0 primary retinal cells (T.L.Y. et al., unpublished data).

Mouse *TUG1* (*mTUG1*) is a novel gene, localized to chromosome 11A1, which is represented by over 200 expressed-sequence tags (ESTs) in Unigene (NCBI, Mm.359796). In order to determine the cDNA structure of *TUG1*, SeqMan (Lasergene) was utilized to assemble EST contigs, which were then aligned to the genomic sequence of the *TUG1* locus (Figure 1A). The *TUG1* ESTs assembled into two major contigs (A and B) consisting of four and two exons, respectively. Contig A is 4.6 kb and encodes a polyadenylation signal (AATAAA) 20 nucleotides from the predicted 3′ end, with multiple ESTs containing a poly-A tract at the indicated 3′ termination site. Contig B contains 3.6 kb of sequence.

However, after 2.6 kb, many ESTs terminate and contain poly-A tracts. In addition, a polyadenylation signal (AATAAA) is present in the DNA sequence 13 nucleotides upstream of this poly-A tract. Some ESTs continue beyond 2.6 kb, and these ESTs overlap with the 5′ end of exon 4 of *TUG1a*, suggesting that perhaps exon 2 of contig B may extend further in the 3′ direction to the same terminal location as contig A. This would lead to a prediction of putative cDNAs for contig B of 2.6 and 6.7 kb.

Northern blots were performed with RNA from retinas of different developmental time points and from multiple adult tissues in order to confirm the size of the *TUG1* RNAs (Figure 2A). *TUG1* was found to be expressed throughout retinal development, with a peak of expression between E18 and P6 (Figure 2A). This peak of expression was confirmed by quantitative RT-PCR analysis (data not shown). The major band was between 6 and 9 kb, and the minor band between 4 and 5 kb, consistent with the predicted sizes of *TUG1a* (4.7 kb) and *TUG1c* (6.7 kb). Both cDNAs also were present in adult tissues (Figure 2B). *TUG1* appears to be expressed in multiple adult-mouse tissues including brain, heart, kidney, and, to a lesser extent, muscle and spleen. Thus, *TUG1a* and *TUG1c* are the major *TUG1* transcripts present during development in the retina and in the mature tissues examined. *TUG1b*, which was predicted based on the presence of a putative polyadenylation site at its 3′ end in several ESTs, may exist at a lower level in these tissues or in other tissue types.

A BLAST (NCBI) search with *mTUG1* nucleotide sequence revealed multiple mouse ESTs (all representing the same locus, described above). The 3′-most 3 kb segment of *TUG1* ESTs was identical to the predicted 5′ end of a gene fragment called TI-227. TI-227 was identified in a screen for RNAs that are more highly expressed in a melanoma line that metastasizes to the lung than in a line that does not [16]. TI-227 was described (by Northern-blot analysis) as a 3.5 kb transcript, present in a metastatic melanoma line that was not expressed in "normal" tissues [17]. On the Basis of the studies presented here, either the analysis of TI-227 failed to correctly determine the location of the 5′ end and to detect its normal expression, or there exists another gene or *TUG1* isoform that is not expressed in wild-type cells and that overlaps the *mTUG1* locus and extends beyond its 3′ end. Other than TI-227, no other homology for *mTUG1* to any known mouse gene was found. Highly conserved homologs were identified in rats, dogs, cows, and humans, but no homolog was found in other organisms (e.g., *C. elegans*, *D. melanogaster*, *X. laevis*, *D. rerio*, *F. rubripes*, and *G. gallus*) in searches of the Genbank database.

Contigs were assembled of ESTs from the human *TUG1* (*hTUG1*) Unigene cluster via SeqMan (Lasergene) in order to predict the *hTUG1* cDNA sequence (Figure 1B). On the basis of the number (over 200) and origin of the ESTs, human *TUG1* appears to be expressed at a very high level and in a variety of tissues (NCBI, Hs.158783). The predicted cDNAs contain an exonic structure very similar to that of *mTUG1a*. However, *mTUG1a* appears to contain an additional 57-nucleotide exon (exon 1). Although this sequence was not present in any of the *hTUG1* ESTs, the identical 57-nucleotide sequence was present 5′ of *hTUG1* contig A,
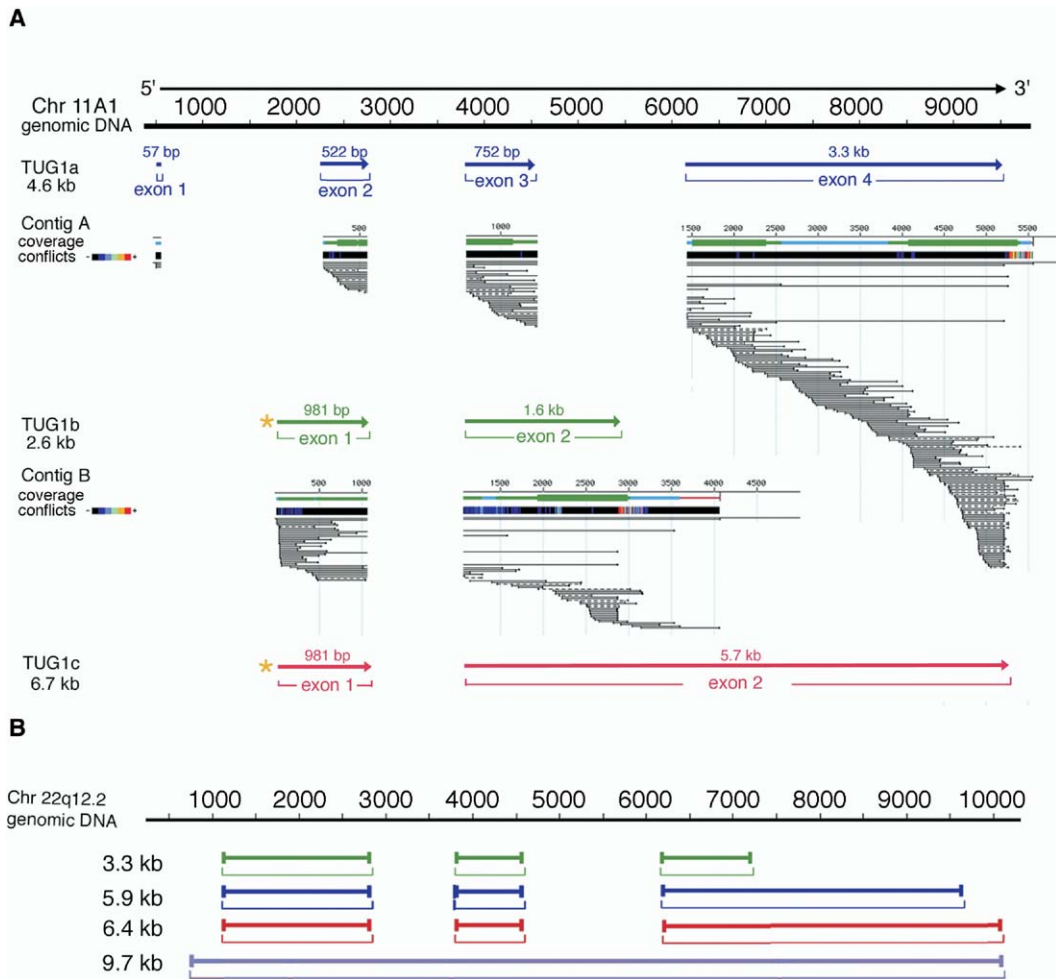
Figure 1. Genomic Organization of *TUG1*

(A) Murine *TUG1* was localized to chromosome 11A1. Its genomic sequence was obtained from NCBI. ESTs for *mTUG1* were downloaded from Unigene and assembled into contigs with SeqMan (Lasergene). Two major contigs that represent two presumptive cDNAs were assembled: *TUG1a* and *TUG1b*. Visual inspection of the contigs revealed that exon 4 of *TUG1a* had considerable overlap with exon 2 of *TUG1b*, suggesting that a third cDNA, *TUG1c*, may exist in vivo. Thin black lines show locations of the ESTs in relation to the genomic locus and putative cDNAs. Solid black lines represent ESTs originally sequenced in the 5′-to-3′ direction and dashed lines represent ESTs sequenced in the 3′-to-5′ direction. For each contig, the conflict between the ESTs is represented by a colored bar beneath the presumptive cDNAs (color scale shown). Orange asterisks in *TUG1b* and *TUG1c* represent the location of a genomic poly-T tract consisting of approximately 30 thymidine residues.

(B) Human *TUG1* was localized to chromosome 22, and the genomic sequence for this locus was obtained from NCBI. ESTs for *hTUG1* were downloaded from Unigene and assembled into contigs with SeqMan (Lasergene). Five major contigs were assembled. Visual inspection of the contigs and polyadenylation sites suggested the presence of four presumptive cDNAs.

suggesting that this 57-nucleotide sequence (*TUG1a*, nt 1–57) may represent an important regulatory region.

### Expression of an Opposite Strand Transcript at the *TUG1* Locus

The presence of terminal poly-A tracts in numerous *TUG1* ESTs indicates that the most common isoforms of *TUG1* RNA run in the 5′-to-3′ direction, as diagrammed in Figure 1A, and are referred to here as the "+" strand. In the human and mouse genomes, greater than 98% of splice-donor and -acceptor sites are GT-AG, and the majority of those remaining are GC-AG [18, 19]. Therefore, one can predict the directionality of a cDNA on the basis of the intronic sequence that contains the splice donor and acceptor. All three of the *mTUG1* splice junctions have the conserved GT-AG

splice-donor and -acceptor sites in the 5′-to-3′ direction denoted in Figure 1.

Multiple ESTs terminate at the same 5′-end location indicated for *TUG1b* and *TUG1c* in Figure 1. At this genomic location, there exists a string of approximately 30 thymidine residues (see Figure 1A, orange asterisks). The common use of poly-T primers for generation of ESTs by reverse transcription may explain why so many ESTs have an end at this location. However, in order for priming with a poly-T primer to occur here, an RNA transcribed from the opposite strand (a "−" strand transcript) that contained the transcribed-polyadenosine tract would have to have been produced. In order to examine whether such a transcript is made, Northern-blot hybridization analyses were performed with a strand-specific probe that would recog-
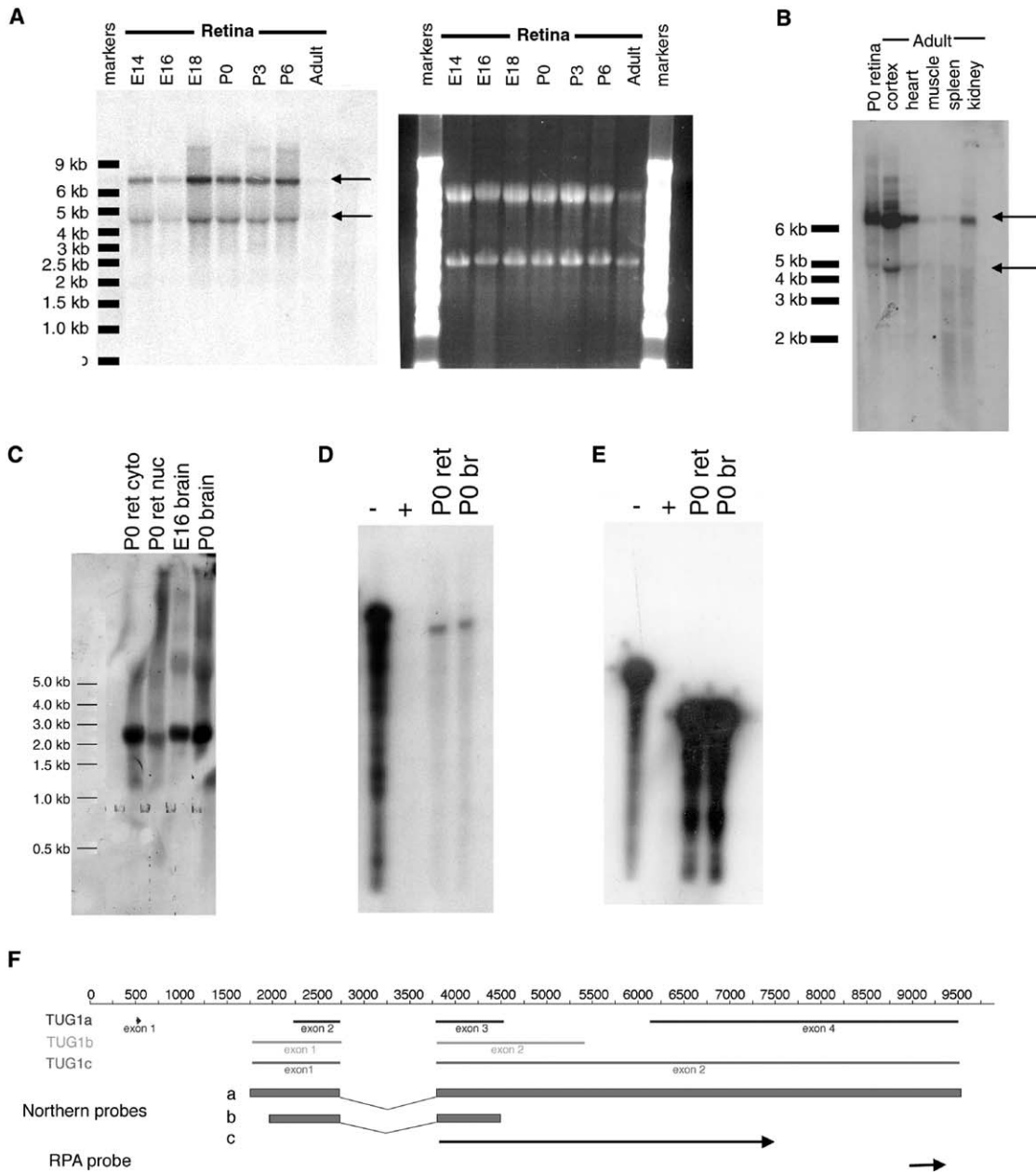
Figure 2. Northern-Blot Analyses of *mTUG1* in the Developing Retina and in Adult Tissues

(A) RNA was harvested from murine retinas of several developmental stages and used for Northern blots. Ethidium-bromide staining of the RNA reveals 28S and 18S ribosomal RNAs in order to provide an estimate of relative RNA levels among samples. A double-stranded probe that recognizes all presumptive *TUG1* cDNAs was used. Arrows mark bands of approximately 6.7 kb and 4.7 kb.

(B) RNA from various adult tissues was harvested, and Northern blots were performed with a full-length double-stranded *TUG1c* probe. Arrows mark bands of approximately 6.7 kb and 4.7 kb.

(C) Northern blot probed with a single-stranded-RNA probe that recognizes a transcript that is transcribed in the – direction, opposite to *TUG1a–TUG1c*. This transcript is approximately 2.5 kb and is found in both cytoplasmic and nuclear fractions of P0 retinas and in total E16 and P0 brain.

(D and E) RNase protection assays (RPAs) with strand-specific probes that recognize *TUG1* RNA transcribed in the – (D) and β-actin (E). Each set of RPAs consists of the following: –, yeast target RNA (tRNA) without addition of RNase (free probe); +, yeast tRNA with RNase (negative control); and either P0 retina or brain total RNA as tRNA with addition of RNase.

(F) Schematic of probe locations in relation to *TUG1* cDNAs. Northern-blot probes a and b are DNA probes that recognize *TUG1* cDNAs in (A) and (B), respectively. Northern-probe c is an RNA probe used in (C).

nize – strand RNAs. This probe recognized a band of approximately 2.5 kb in P0 retina and E16 and P0 brain (Figure 2C). The 2.5 kb transcript was not observed with

a full-length, nick-translated double-stranded-*TUG1* probe or a 5′ *TUG1* double-stranded probe, as shown in Figures 2A and 2B, indicating that the – strand tran-

script at the *TUG1* locus is expressed at a much lower level than the + strand transcripts, *TUG1a* and *TUG1c*.

Ribonuclease (RNase) protection assays were performed with a + strand *TUG1* probe (Figure 2D) and a β-actin probe (Figure 2E) in order to confirm the presence of a – strand transcript. The probe to *TUG1* protected an RNA fragment that was present in both P0 retina and brain samples. These data suggest the presence of a *TUG1* RNA transcribed from the opposite strand to that of *TUG1a–c*.

### *TUG1* as a Noncoding RNA

Full-length (6.7 kb) *TUG1c* was assembled by amplifying different regions of *TUG1* via PCR and ligating them with two ESTs that covered the extreme 5′ and 3′ ends. Multiple clones were sequenced to confirm the predicted *TUG1* sequence. *TUG1* sequence was consistent between the sequencing of our amplified clones, the contigs of ESTs, and the genomic sequence reported by NCBI.

Within each of the predicted cDNAs (*mTUG1a–c*), no ORF encoding more than 100 amino acids was observed. ORFs greater than 50 amino acids in both orientations are shown in Figure 3A. Alignment of mouse and human putative full-length cDNAs revealed a high level of conservation between the *TUG1* homologs in multiple regions of the cDNAs (Figure 3A). However, only one of the mouse ORFs, Mouse ORF1, is in the same reading frame as its corresponding human ORF. This ORF is the best candidate for a true coding region, with 40 of its 51 amino acids being identical between mouse and human (Figure 3C). Mouse ORF1 is encoded only by *TUG1b* and *TUG1c*; *TUG1a*, a major transcript in most tissues (Figure 2B), splices around the ORF sequence and would therefore not encode the putative peptide. *TUG1* homologs were not identified in the incomplete genomes for chimpanzees, cats, sheep, or pigs. *TUG1* loci were identified in rats, dogs, and cows, and these homologs were over 85% identical in the region of ORF1. However, in the cow, a homologous ORF to ORF1 was identified but terminated prematurely after only 31 amino acids (Figure 3C). In both rats and dogs, a homologous start methionine was not found in the vicinity of the identified *TUG1* locus.

In vitro transcription and translation experiments performed with *TUG1c* failed to produce any protein product (data not shown).

### *TUG1* Is Expressed in Multiple Regions of the CNS

In situ hybridization analyses were performed with *TUG1* probes in order to investigate more specifically where *TUG1* is expressed in the CNS. These studies indicated that *TUG1* is expressed within the developing retina from E14 to P0 in retinal regions where progenitor cells reside (Figures 4A–4C). However, as the retina matures, expression of *TUG1* becomes most highly expressed in the inner nuclear layer and in the ganglion-cell layer (Figure 4D). *TUG1* expression also was analyzed in the P0 mouse nervous system. Figure 4 shows that *TUG1* is expressed at P0 in multiple regions, including the spinal cord (E), olfactory epithelium (F), hippocampus (G), and cortex (H).

### *TUG1* Is Required for Normal Photoreceptor Development

Gain- and loss-of-function experiments were performed in order to investigate the function of *TUG1* in the developing retina. A full-length *TUG1c* cDNA driven by a ubiquitous promoter (CAG) was coelectroporated with CAG-GFP in vivo into the P0 rat retina [20]. The coelectroporation rate has been found to be over 85% under the conditions applied here [20]. Retinas were harvested, fixed, and sectioned at P14. No obvious phenotype was observed among the electroporated cells; this might have been predicted given the relatively high endogenous levels of *TUG1* in the retina at this time. An RNAi construct was thus constructed to determine whether a reduction in *TUG1* would reveal a requirement for *TUG1*.

A small interfering RNA (siRNA) hairpin that targets exon 2 of *TUG1a* and exon 1 of *TUG1b* and *TUG1c* (*TUG1*-RNAia, see Experimental Procedures) was driven by the RNA-polymerase-III-driven U6 promoter [21]. This construct was electroporated in vivo into the mouse retina at P0. Membrane bound green fluorescent protein (mGFP) driven by the CAG promoter was coelectroporated in order to visualize which cells were electroporated. As a control for nonspecific effects of electroporation and activation of the RNAi machinery, retinas were coelectroporated with a U6 vector targeting GAPDH and with CAG-mGFP. Retinas were collected and sectioned at P25, and GFP was visualized in order to examine the morphology of transfected cells. Results of these experiments are shown in Figure 5. As previously described, retinal cells electroporated in vivo at P0 via this method are primarily mitotic progenitor cells or newly postmitotic cells [20]. The majority (over 70%) of these electroporated cells ultimately become rod-photoreceptor cells. This is true for coelectroporation of CAG-GFP and an empty U6 vector [20] as well as for coelectroporation of CAG-GFP with an RNAi targeting GAPDH (Figures 5A and 5E). This RNAi construct for GAPDH has been shown to effectively knock down RNA levels of GAPDH but to have no cellular phenotype in the retina [20]. Similar to control electroporations, the majority of retinal cells electroporated with *TUG1* RNAi resided in the outer nuclear layer, where photoreceptor cells reside. However, unlike control electroporated cells, which extended long, rod-like outer segments (Figure 5A), *TUG1* RNAi-electroporated cells appeared abnormal. Many of the photoreceptors completely lacked inner or outer segments (Figure 5B). In addition, some of the cell bodies electroporated with *TUG1* RNAi were present in the inner nuclear layer or inner plexiform layer but, similar to photoreceptor cells, extended processes through the outer nuclear layer (Figure 5B). In addition to the results shown for the *TUG1* RNAi construct shown in Figure 5, a second RNAi construct targeting *TUG1* (*TUG1* RNAib, see Experimental Procedures) also resulted in the loss or shortening of outer segments, although with a lower penetrance.

The experiment described above was performed by coelectroporating either *TUG1* RNAi or GAPDH RNAi with CAG-mGFP in order to molecularly analyze which cell types were generated after *TUG1* RNAi. Instead of being sectioned at P21, the electroporated retinas were
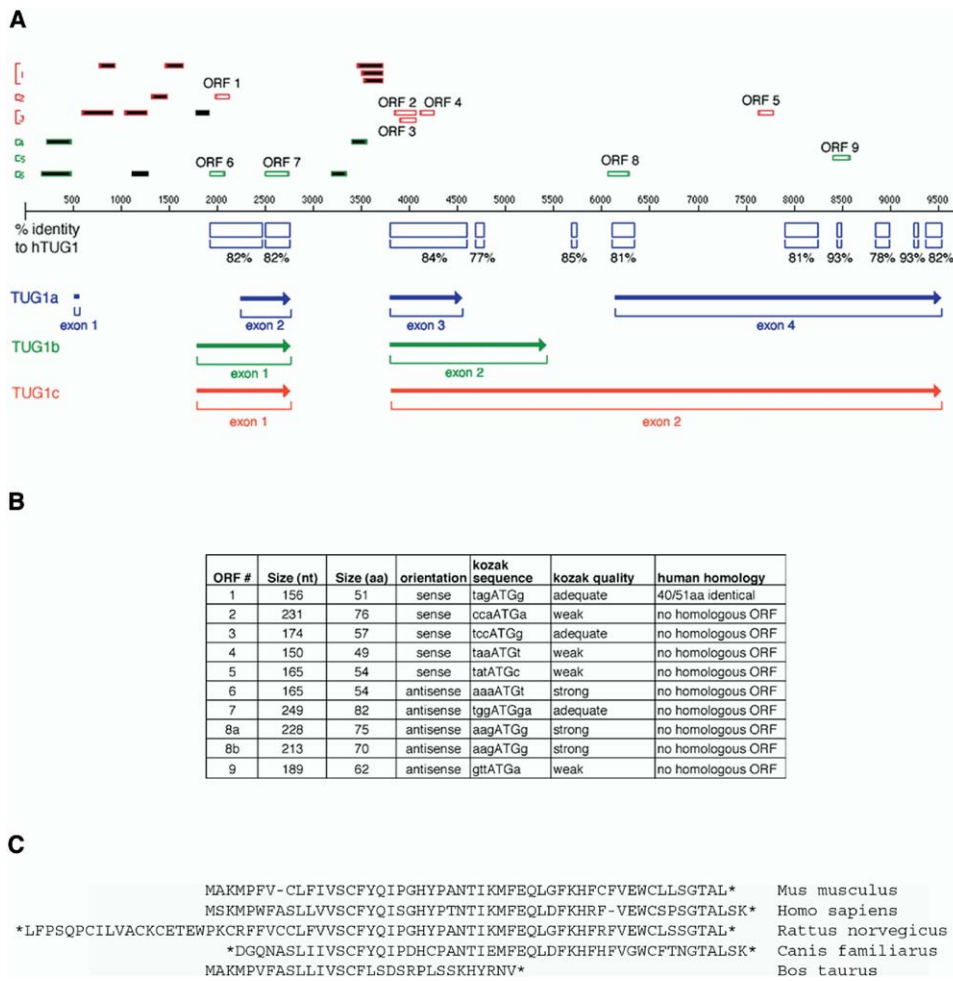
Figure 3. Possible Open Reading Frames at the *mTUG1* Locus

(A) Relative locations of open reading frames (ORFs) greater than 50 amino acids in length in all six reading frames in relation to the *mTUG1* genomic locus and presumptive cDNAs. ORFs in the + strand are represented by red boxes, and those in the − strand are outlined in green. Open boxes are those ORFs that overlap with *TUG1* cDNAs. Open blue boxes represent regions of highest homology between *mTUG1* and *hTUG1*, with percentage of homology listed.

(B) Table describing the characteristics of the nine ORFs greater than 50 amino acids in length present at the murine *TUG1* locus. Kozak quality was determined based on guidelines previously described [41].

(C) Alignment of *mTUG1* ORF1 with the homologous *TUG1* ORFs in rats (*Rattus norvegicus*), humans (*Homo sapiens*), cows (*Bos taurus*), and dogs (*Canis familiaris*). Asterisks denote stop codons.

dissociated, plated, and stained with antibodies that mark different retinal cell types. After *TUG1* RNAi transfection, the majority of retinal cells expressing GFP also expressed rhodopsin. However, a significant decrease of approximately 10% in the percentage of rhodopsin-positive cells that were also GFP positive was observed (Figure 5E). With this decrease, a concomitant increase to approximately 10% in the percentage of transfected cells expressing peanut agglutinin (PNA), a cone-photoreceptor marker, was observed (Figures 5C–5E). No significant change was observed in the percentage of cells expressing Chx10 (which marks bipolar cells), Pax6 (which marks amacrine cells), or glutamine synthetase (which marks Müller glial cells and undifferentiated cells).

The effects of *TUG1* knock down 3 days after in vivo electroporation of the construct were analyzed in order to begin to study the molecular events that lead to the production of cells with cone properties. In a manner similar to that of the experiments described above, CAG-mGFP was coelectroporated either with the RNAi construct targeting *TUG1* or with an empty U6 vector. After three days, retinas were harvested, dissociated, and then sorted with fluorescence-activated cell sorting (FACs) to collect GFP-positive electroporated cells. For each independent experiment, 3–4 retinas were pooled and approximately 500,000–1,000,000 cells were collected. From these cells, RNA was purified and cDNA was synthesized and amplified. Amplified cDNA was directly labeled with Cy3 or Cy5 and hybridized to cDNA microarrays containing approximately 12,000 brain and retinal cDNAs, isolated by Dr. Bento Soares, and 400 lab clones. This experiment was performed two independent times, and Cy dyes were inverted for both experiments to yield four data points for each variable. *TUG1* was found by microarray analysis to be
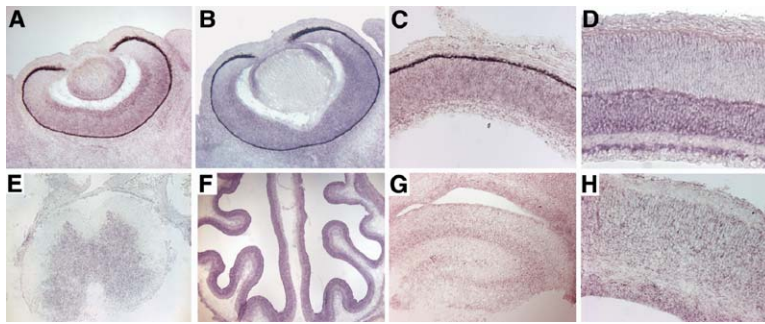
Figure 4. In Situ Hybridization Analysis of *mTUG1* in the Developing Nervous System

(A–D) In situ hybridizations with *TUG1* specific probes, a – strand probe to recognize *TUG1a–c* (see Experimental Procedures) on retinal sections at E14 (A), E16 (B), P0 (C), and P10 (D) showing *TUG1* expression in retinal progenitor and precursor cells (A–C) and in inner nuclear layer and ganglion-layer cells (D).

(E–H) In situ hybridizations with *TUG1*-specific probes. The hybridizations show *TUG1* expression at P0 in spinal cord (E), olfactory epithelium (F), hippocampus (G), and cortex (H).

downregulated between 5.3-fold and 5.9-fold after *TUG1* RNAi electroporation. In addition, a number of other genes were found to be changed reproducibly. Table 1 shows a partial list of the genes found to be consistently upregulated (upper section of Table 1) or downregulated (lower section of Table 1) upon *TUG1* RNAi transfection. Cone arrestin was reproducibly upregulated after *TUG1* knock down, and this is consis-

tent with the hypothesis that *TUG1* is involved in regulating photoreceptor-gene expression. However, rod genes including rod outer-segment-membrane protein 1 (ROM1), phosphodiesterase 6B (Pde6b), and cyclic-nucleotide-gated channel $\alpha$1 (Cnga1) also were upregulated (upper section of Table 1). In addition, two genes that were consistently downregulated with *TUG1* knockdown were the transcription factors cone-rod
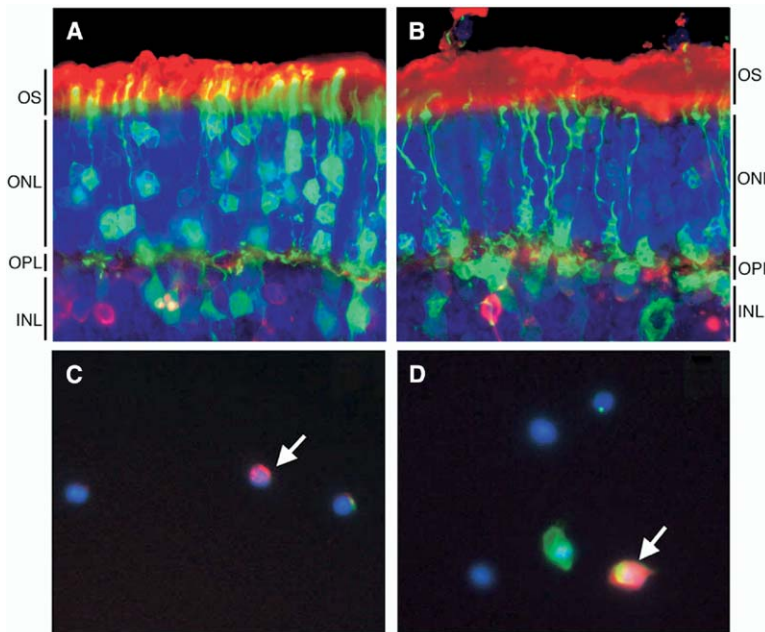


Figure 5. Knockdown of *TUG1* from P0 by RNAi Constructs Targeting *TUG1* in the Retina

(A and B) A construct encoding GFP was coelectroporated in vivo into P0 retinas along with a control hairpin targeting GAPDH (A) or a *TUG1* RNAi construct (B). Retinas were harvested at P25 and sectioned and stained for DAPI (blue). Electroporated cells were visualized by GFP fluorescence (green). Rod-photoreceptor outer segments were immunostained with anti-rhodopsin (red). At least three independent retinas were analyzed for each variable. The following abbreviations are used: OS, outer segments; IS, inner segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; and GCL, ganglion-cell layer.

(C–E) P0 retinas were coelectroporated with GFP and a GAPDH RNAi U6 vector or with a *TUG1* RNAi U6 vector. Retinas were harvested at P21, dissociated, and immunostained for cell-type-specific markers. Examples are shown of PNA staining (red) of dissociated cells in GAPDH RNAi-electroporated cells (C) and in *TUG1* RNAi cells (D). Electroporated cells are marked with GFP (green), and nuclei are stained with DAPI (blue). Arrows indicate a PNA+GFP– cell in (C) and a PNA+GFP+ cell in (D). GFP-positive cells were counted to analyze the percentage of electroporated cells that expressed markers of different cell types (E). Three independent retinas were analyzed for each combination. Error bars represent the standard deviation for three independent retinas. Significance tests comparing control GAPDH U6-electroporated cells with TUG RNAi-electroporated cells were performed. * $p < 0.05$; ** $p < 0.001$.
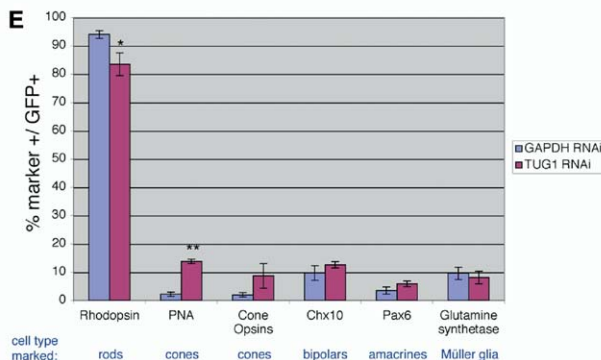
Table 1. Genes Changed after Knockdown of *TUG1*

| Genbank ID | Unigene ID | Name | arr1 | arr2 | arr3 | arr4 | ave | stdev |
|---|---|---|---|---|---|---|---|---|
| BE950239 | Mm.347492 | rod outer segment membrane protein 1 | 2.82 | 2.22 | 0.73 | 1.64 | 1.85 | 0.77 |
| 31247A_I13 | Mm.1372 | Phosphodiesterase 6B, cGMP, rod receptor, beta | | 1.41 | 1.56 | 2.27 | 1.75 | 0.38 |
| 31247A_J10 | Mm.23793 | cyclic nucleotide gated channel, alpha 1 | | 2.44 | 1.1 | 3.57 | 2.37 | 1.01 |
| 31247A_J09 | | cone arrestin | 1.09 | 4.55 | 1.52 | 2.08 | 2.31 | 1.34 |
| BE953626 | Mm.240434 | programmed cell death 8 (AIF) | 2.72 | 3.03 | 0.94 | 2.13 | 2.20 | 0.80 |
| BF461723 | Mm.29816 | programmed cell death 6 interacting protein (Alix/AIP1) | 2.61 | 1.35 | | 2.44 | 2.13 | 0.56 |
| BE988698 | Mm.2159 | BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3 | 3.03 | 3.33 | 0.83 | 1.96 | 2.29 | 0.98 |
| BE981505 | Mm.2159 | BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3 | 1.65 | 1.79 | 0.94 | 2.22 | 1.65 | 0.46 |
| BE995440 | Mm.147091 | CUG triplet repeat,RNA binding protein 2, NAPOR | 2.3 | 2.44 | 0.73 | | 1.82 | 0.77 |
| BE983994 | Mm.147091 | CUG triplet repeat,RNA binding protein 2, NAPOR | | 2.22 | 0.73 | 2.17 | 1.71 | 0.69 |
| BF467175 | Mm.5027 | enhancer of zeste homolog 1 (Drosophila) | | 0.29 | 1.78 | 2.04 | 1.37 | 0.77 |
| BE985513 | Mm.46766 | RIKEN cDNA 2610528E23 gene | 2.57 | 2.13 | | 7.14 | 3.95 | 2.27 |
| BF466153 | Mm.46766 | RIKEN cDNA 2610528E23 gene | 2.48 | | 2.18 | 1.32 | 1.99 | 0.49 |
| BF463356 | Mm.46513 | RIKEN cDNA 2610200G18 gene | 1.64 | 2.70 | | 1.82 | 2.05 | 0.46 |
| BE982942 | Mm.46513 | RIKEN cDNA 2610200G18 gene | 2.68 | 1.79 | 0.82 | 1.79 | 1.77 | 0.66 |
| AI852207 | Mm.4541 | SRY-box containing gene 2 | 2.03 | 3.85 | 0.71 | 3.57 | 2.54 | 1.26 |
| BE954949 | Mm.3304 | neuron specific gene family member 2 | 2.45 | 1.67 | 0.89 | 2.08 | 1.77 | 0.58 |
| BE988688 | Mm.3304 | neuron specific gene family member 2 | 1.54 | 2.50 | 0.77 | 1.52 | 1.58 | 0.61 |
| BF455644 | Mm.329322 | formin-family protein FHOS2 | | 3.03 | 1.04 | 2.33 | 2.13 | 0.82 |
| BE987275 | Mm.329322 | formin-family protein FHOS2 | 2.81 | 1.69 | 0.78 | 1.43 | 1.68 | 0.73 |
| BF466074 | Mm.31646 | actin-like 6 | | 6.25 | 0.83 | 2.13 | 3.07 | 2.31 |
| BF467587 | Mm.31646 | actin-like 6 | 3.04 | 1.67 | 1.45 | 3.03 | 2.30 | 0.74 |
| BF460725 | Mm.29939 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9 | 2.45 | 5.26 | 1.07 | 3.45 | 3.06 | 1.53 |
| BF467681 | Mm.29939 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9 | 3.3 | 1.61 | 1.96 | 1.85 | 2.18 | 0.66 |
| BE985600 | Mm.29867 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2 | 2.46 | 1.69 | 0.75 | 1.47 | 1.59 | 0.61 |
| BE954716 | Mm.291928 | catenin beta | 2.79 | 4.55 | | 2.13 | 3.15 | 1.02 |
| BF463481 | Mm.291624 | anaphase promoting complex subunit 2 | 5.09 | 2.33 | 1.05 | 2.50 | 2.74 | 1.47 |
| BE986838 | Mm.290995 | collapsin response mediator protein 1 | 1.4 | 3.03 | 0.95 | 2.38 | 1.94 | 0.81 |
| BE952133 | Mm.290610 | Notch gene homolog 1 (Drosophila) | 1.75 | 3.85 | 0.62 | 1.85 | 2.02 | 1.16 |
| AW048812 | Mm.280029 | hairy and enhancer of split 6 (Drosophila) | 1.49 | 3.57 | 0.53 | 3.45 | 2.26 | 1.30 |
| BE987256 | Mm.277099 | Neuroligin 1 (Nlgn1), mRNA | 2.52 | 1.89 | 0.93 | 1.45 | 1.70 | 0.58 |
| BI202386 | Mm.259222 | nudix (nucleoside diphosphate linked moiety X)-type motif 8 | 1.88 | | 1.37 | 2.04 | 1.76 | 0.29 |
| BE995150 | Mm.259103 | polyhomeotic-like 2 (Drosophila) | 1.78 | | 1.37 | 2.00 | 1.72 | 0.26 |
| BE955737 | Mm.258771 | protein tyrosine phosphatase, receptor type, S | 2.18 | 1.56 | | 2.70 | 2.15 | 0.47 |
| BF464976 | Mm.239605 | B-cell translocation gene 2, anti-proliferative | 3.52 | 5.88 | 0.79 | 1.49 | 2.92 | 1.98 |
| BE949664 | Mm.14802 | H19 fetal liver mRNA | | 1.49 | 0.89 | 2.33 | 1.57 | 0.59 |
| 31247A_C5 | | p16 | | 2.22 | 0.75 | 2.63 | 1.87 | 0.81 |
| 31247A_G13 | | CDK4 | | 2.13 | 1.42 | 2.04 | 1.86 | 0.32 |

| Genbank ID | Unigene ID | Name | arr1 | arr2 | arr3 | arr4 | ave | stdev |
|---|---|---|---|---|---|---|---|---|
| 31244A_A4 | Mm.134516 | Otx2 | 0.31 | 1 | 0.2 | 0.478 | 0.497 | 0.307 |
| 31247A_I23 | Mm.134516 | Otx2 | 0.17 | 1.408 | 0.22 | 0.33 | 0.532 | 0.509 |
| 31244A_B9 | Mm.134516 | Otx2 | 0.53 | 1.299 | 0.19 | 0.324 | 0.586 | 0.429 |
| 31247A_I10 | Mm.134516 | Otx2 | 0.16 | 1.818 | 0.44 | 0.437 | 0.714 | 0.648 |
| 31246A_G2 | | Crx | 0.38 | 1.786 | 0.2 | 0.362 | 0.682 | 0.641 |
| 31247A_H09 | | Crx | 0.28 | 1.852 | 0.27 | 0.37 | 0.693 | 0.67 |
| 31247A_H20 | | Crx | 0.24 | 2.174 | 0.26 | 0.508 | 0.795 | 0.803 |
| BE950590 | Mm.9001 | nemo like kinase | 0.67 | 1.124 | 0.17 | 0.189 | 0.538 | 0.393 |
| 31246A_D6 | | Math3 | 0.56 | 1.818 | 0.49 | 1.25 | 1.03 | 0.544 |
| 31247A_I01 | | Dkk3 | 0.73 | 0.513 | 0.32 | 0.855 | 0.604 | 0.205 |
| BE953314 | Mm.119162 | ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 | 0.23 | 0.242 | 0.22 | 0.353 | 0.261 | 0.054 |
| BF461840 | Mm.119162 | ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 | 0.3 | 0.346 | 0.42 | 1.042 | 0.527 | 0.3 |
| BF464254 | Mm.132238 | CREB binding protein | 1.32 | 0.431 | 0.6 | 0.73 | 0.77 | 0.335 |
| BE950538 | Mm.144699 | nudix (nucleotide diphosphate linked moiety X)-type motif 3 | 0.34 | 0.379 | 0.62 | 1.667 | 0.751 | 0.539 |
| BE951272 | Mm.144699 | nudix (nucleotide diphosphate linked moiety X)-type motif 3 | 0.42 | 0.735 | 0.74 | 1.724 | 0.905 | 0.49 |
| BE953493 | Mm.197552 | transforming growth factor, beta receptor I | 0.78 | 0.571 | 0.41 | 0.719 | 0.62 | 0.143 |
| BE988861 | Mm.220038 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 5 | 0.22 | 0.233 | 0.63 | 1.176 | 0.565 | 0.39 |
| BE949675 | Mm.220038 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 5 | 0.29 | 0.251 | 0.74 | 1.299 | 0.645 | 0.424 |
| BF460946 | Mm.236513 | poly(rC) binding protein 2 | 0.48 | 0.513 | 0.15 | 0.264 | 0.352 | 0.151 |
| BE950149 | Mm.236513 | poly(rC) binding protein 2 | 0.47 | 0.763 | 0.17 | 0.25 | 0.413 | 0.23 |
| BE951954 | Mm.275521 | Wilms' tumour 1-associating protein | 0.6 | 0.379 | 0.71 | 0.787 | 0.619 | 0.154 |
| BE983748 | Mm.4502 | minichromosome maintenance deficient 3 (S. cerevisiae) | 0.59 | 0.427 | 0.48 | 0.775 | 0.568 | 0.133 |
| BE988418 | Mm.28123 | Fas-activated serine/threonine kinase | 0.48 | 0.962 | 0.16 | 0.435 | 0.509 | 0.289 |
| BE952305 | Mm.298 | solute carrier family 25, member 3 | 0.35 | 0.388 | 0.61 | 0.826 | 0.544 | 0.191 |
| BE995394 | Mm.30059 | myristoylated alanine rich protein kinase C substrate | 0.47 | 0.457 | 0.21 | 0.481 | 0.404 | 0.113 |
| BE996335 | Mm.30059 | myristoylated alanine rich protein kinase C substrate | 0.96 | 0.297 | 0.56 | 0.709 | 0.631 | 0.24 |
| AI846170 | Mm.34701 | pumilio 1 (Drosophila) | 0.54 | 2.5 | 0.46 | 1.235 | 1.184 | 0.818 |
| AI848408 | Mm.34871 | inhibitor of DNA binding 2 | 0.19 | 0.699 | 0.74 | 2.5 | 1.032 | 0.875 |
| BE950254 | Mm.3918 | ATP-binding cassette, sub-family A (ABC1), member 4 | 0.54 | 0.621 | 0.47 | 0.775 | 0.602 | 0.114 |
| BE950117 | Mm.41665 | glutamate receptor, ionotropic, N-methyl D-asparate-associate | 0.69 | 0.388 | 0.16 | 0.141 | 0.345 | 0.222 |
| BF461223 | Mm.4405 | C. elegans ceh-10 homeo domain containing homolog | 0.65 | 0.962 | 0.35 | 1.299 | 0.815 | 0.353 |
| BE985957 | Mm.4657 | amyloid beta (A4) precursor protein-binding, family A, member | 0.4 | 0.35 | 0.74 | 2.083 | 0.893 | 0.703 |
| BE954299 | Mm.316975 | Similar to 40S ribosomal protein S20 (LOC383912), mRNA | 0.47 | 0.369 | 0.66 | 1.695 | 0.798 | 0.528 |
| BF462127 | Mm.320246 | Transcribed sequences | 0.22 | 0.397 | 0.62 | 0.503 | 0.435 | 0.147 |
| BE985546 | Mm.328963 | zinc finger protein 297 | 0.5 | 0.45 | 0.61 | 1.136 | 0.674 | 0.273 |
| BF460670 | Mm.332303 | protein tyrosine phosphatase, receptor type, K | 0.34 | | 0.15 | 0.272 | 0.254 | 0.079 |
| BE953195 | Mm.332522 | WD repeat and FYVE domain containing 3 | 0.53 | 0.386 | 0.14 | 0.214 | 0.318 | 0.152 |
| BE950505 | Mm.345876 | rcd1 (required for cell differentiation) homolog 1 (S. pombe) | 0.74 | 0.917 | 0.41 | 0.347 | 0.604 | 0.235 |

P0 retinas were coelectroporated in vitro with a construct encoding GFP along with an empty U6 vector or a U6 RNAi construct targeting *TUG1*. Retinas were harvested at P3 and dissociated. Retinal cells were sorted by fluorescence-activated cell sorting (FACs), and GFP-positive cells were collected. From these samples, GFP-positive cells were lysed, RNA was purified, and cDNA was synthesized. cDNA samples were amplified and labeled with Cy3 or Cy5 and hybridized to a cDNA microarray. Two independent experiments comparing TUG RNAi-electroporated cells (Cy5) to U6-empty-vector-electroporated cells (Cy3) were performed (arr1 and arr3). Microarrays 2 and 4 (arr2 and arr4) are replicate hybridizations of arr1 and arr3 with the fluorescent labels reversed; the inverse values are shown. Numbers represent the multiple of the change over the baseline of expression of the indicated gene upon *TUG1* RNAi transfection. Genes highlighted in yellow are photoreceptor genes; genes in blue are cell-death-linked genes.

The upper section of this table shows genes upregulated after knockdown of *TUG1*, and the lower section of this table shows genes downregulated after knockdown of *TUG1*.
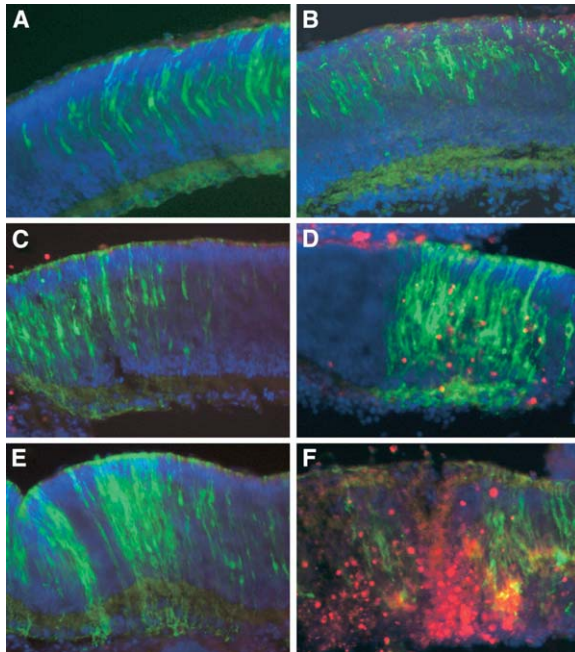
Figure 6. *TUG1* Knockdown Results in an Increase in TUNEL-Positive Cells

P0 retinas were coelectroporated in vivo with constructs encoding GFP and an RNAi construct targeting either GAPDH (A, C, and E) or *TUG1* (B, D, and F). At P21, retinas were dissected, fixed, and sectioned. Electroporated cells were visualized by GFP, and sections were stained with TUNEL to mark dying cells. In three independent retinas, no TUNEL-positive cells were observed in GAPDH-electroporated regions. In *TUG1* RNAi-electroporated retinas, varying numbers, ranging from none (B) to moderate numbers (D) to high numbers (F), of TUNEL-positive cells were observed.

homeobox protein (Crx) and orthodenticle homolog 2 (Otx2) (lower section of Table 1).

### *TUG1* Is Involved in Cell Survival

*TUG1* electroporation yielded fewer GFP-positive cells than did control electroporations in a subset of retinal sections, suggesting that electroporation of *TUG1* RNAi can result in cell death. In addition, several genes involved in the cell-death pathway are upregulated 3 days after *TUG1* RNAi electroporation. These upregulated genes include *programmed cell death 8* (*AIF*), *programmed cell death 6 interacting protein* (*Alix/AIP1*), *Bcl2* (*NIP3*), *CUG triplet repeat*, and *RNA-interacting protein 2* (*NAPOR*) (upper section of Table 1, [22–26]). Therefore, a TUNEL assay was performed to determine whether *TUG1* RNAi resulted in an increase in apoptosis. P0 retinas were coelectroporated with CAG-mGFP and either the GAPDH RNAi construct or the *TUG1* RNAi construct. Retinas were harvested after 4 days and fixed, sectioned, and stained for TUNEL (Roche). Transfected cells were visualized by GFP fluorescence. At P4, many of the transfected cells would still be progenitor cells. Retinal progenitor cells have processes that extend throughout the radial dimension of the retina, with the nuclei migrating up and down with the cell cycle (Figure 6). The morphology of cells electroporated with the *TUG1* RNAi construct was similar in most of

the cases to the morphology of control transfected cells at P4 (Figure 6). In three GAPDH RNAi-electroporated retinas, no TUNEL-positive cells were observed within the main body of the retinas (Figures 6A, 6C, and 6E). However, TUNEL-positive cells were observed where the retinas had been damaged (Figure 6C). In contrast, TUNEL-positive cells were observed within the retina after *TUG1* RNAi electroporation (Figures 6D and 6F). The extent of the TUNEL staining varied in *TUG1* RNAi-electroporated regions of three independent retinas. Some regions of electroporation contained no TUNEL-positive cells (Figure 6B), whereas others had scattered positive cells (Figure 6D) or numerous positive cells (Figure 6F).

### Discussion

#### Expression of *TUG1* Is Induced by Depolarization
*TUG1* was identified in a screen for genes upregulated in developing retinal cells upon addition of taurine. Taurine has been shown to activate both glycine and GABA(A) receptors [27–31]. In the developing retina, taurine has been shown to act through glycine and GABA receptors to stimulate rod-photoreceptor development [6]. In the mature nervous system, stimulation of glycine and GABA receptors results in a hyperpolarization of cells expressing these receptors. However, during development, stimulation of glycine or GABA receptors on an immature neuron results in a depolarization of the cell owing to the gradient of chloride present in immature neuronal cells [32–34]. The application of KCl is a standard method used to depolarize cells. Therefore, during development, the common cellular response to the addition of taurine, glycine, GABA, or KCl is an increase in intracellular calcium levels. *TUG1* was induced by all of these factors (T.L.Y. et al., unpublished data), suggesting that *TUG1* expression is upregulated in response to calcium influx. Thus, we speculate that taurine, through glycine and GABA receptors, stimulates *TUG1* expression in the developing retina.

The observation that *TUG1* is expressed in multiple regions of the nervous system during development is interesting in light of the fact that taurine, glycine receptors, and GABA receptors also are present in these regions during development. Taurine is present at high levels throughout the developing brain, and a role for taurine in CNS development has been suggested by nutrition-deprivation studies [4, 35, 36]. In these studies, pregnant cats were fed a diet that lacked taurine but was otherwise normal. Many of the fetuses of the animals with highly reduced taurine levels were aborted, and those that survived to birth had severely malformed brains and retinas. In addition to the retina, endogenous activation of glycine and GABA receptors by taurine in several regions of the developing nervous system may stimulate *TUG1* expression, in which it also may play a role in the development of these tissues in these regions.

#### *TUG1* Plays a Role in Photoreceptor Development
*TUG1* knockdown affected the morphology of photoreceptor cells, with a loss or shortening of outer segments and a thinning of inner segments relative to con-

trol electroporated rod photoreceptors. In rod and cone photoreceptors, outer segments are the specialized structures that contain all of the phototransduction machinery necessary to absorb light and transform it into a chemical signal that is then transmitted to other cells. Rod and cone photoreceptors share many similar components of this signal-transduction pathway, but they also express unique proteins; for example, rhodopsin is the photopigment in rods, and cone opsins are the photopigments in cones. In addition, rod and cone outer segments appear to be morphologically distinct, with cones having shorter and fatter outer segments than those of rods [37]. In control electroporated retinas at P0, very few to no cones are transfected (Figure 5E). With *TUG1* knockdown, approximately 14% of the electroporated cells expressed PNA, a cone-cell-specific marker (Figure 5E). These results, in conjunction with microarray data, suggest that *TUG1* may normally act to promote rod-photoreceptor genesis and inhibit cone-photoreceptor gene expression within rods.

*TUG1* RNAi transfection resulted in an increase in TUNEL-positive cells at P3, which raised the possibility that the photoreceptor phenotype observed after *TUG1* knockdown may be a result of these cells being sick or dying. However, in retinal-degeneration models, there is no evidence that sick or dying rod photoreceptors begin to express cone markers or that they morphologically resemble cones. In addition, when harvested at P14, the same outer-segment phenotype that was observed at P21 was observed (data not shown), suggesting that the P21-transfected cells never developed outer segments.

Analysis of gene-expression changes upon *TUG1* knockdown revealed that the expression of a number of photoreceptor genes was altered. Some markers of differentiated photoreceptors, such as cone arrestin, ROM1, Pde6b, and Cnga1 were upregulated, whereas the transcription factors Crx and Otx2 were significantly downregulated. Crx and Otx2 have been shown to be required for proper rod development. Conditional knockout of Otx2 in rod-precursor cells resulted in an increase in amacrine cells and a loss of rod-photoreceptor cells [38]. In contrast, targeted deletion of Crx did not inhibit photoreceptors from forming but did prevent their proper differentiation, with a defect in formation of outer segments [7]. In these mutant retinas, the development of other cell types did not appear to be affected. Thus, disregulation of rod and cone genes, along with the lack of proper rod outer-segment formation, may be downstream effects of a reduction in Otx2 and Crx in the absence of *TUG1*.

### *TUG1* Appears to Function as a ncRNA

It is estimated that only 2% of the human genome encodes proteins. In addition, it has been estimated that 98% of the transcribed sequence in eukaryotes does not encode protein, with approximately 95% of pre-mRNA transcribed sequence being intronic sequence [18, 39]. Analyses of mammalian genomes led to estimates of approximately 30,000–40,000 genes [18, 39]. However, these estimates are in contrast to the number of Unigene cDNA clusters, which are between 65,000 and 70,000 genes [40]. This discrepancy is most likely

due to the biases and assumptions used to predict genes and may falsely label regions encoding miRNAs and other ncRNAs as being intergenic.

As a result of the nature of our past molecular techniques and biases, very little is known about the abundance or importance of ncRNAs. Biochemical and biotechnological methods have recently identified a large number of miRNA species that are fairly abundant. However, the identification of large (>1 kb) ncRNA transcripts has been limited to the fortuitous discovery of a small number of regulated or functional RNAs. Analysis of the RIKEN collection of 60,770 full-length cDNAs that represent 33,409 transcriptional units (TUs) revealed that approximately one-third of these TUs appeared to be ncRNAs [12]. Less than 1% of the noncoding TUs did not appear to be polyadenylated, indicating that a large number of these may be transcribed by RNA polymerase II. However, only 29% of the ncRNAs identified were spliced, compared to 82% of known protein-coding transcripts [12]. Although ncRNAs appear to be abundant, very few have been studied in enough detail to show a function of the ncRNA. In this study, we have identified and characterized a gene, *TUG1*, that is spliced and polyadenylated, suggesting that it is transcribed by RNA polymerase II. Although *mTUG1* sequence is highly conserved between mouse and human, only a single ORF (ORF1) of 51 amino acids is conserved between the two species, and this peptide contains no homology to any known protein. This ORF is in a favorable context for translation because it contains an adequate Kozak sequence and is the first ORF from the 5′ end of *TUG1b* and *TUG1c* [41]. It is therefore possible that this ORF produces a 51 amino acid peptide in vivo although it failed to do so in vitro. However, this ORF is only contained within *TUG1b* and *TUG1c* and not in *TUG1a*. There are very few examples of peptides this small being coded directly by a short ORF. Rather, the majority of small peptides are processed from larger proteins. There are limited exceptions to this, with the smallest known ORF encoding a peptide being L41, a ribosomal protein of 25 amino acids [42]. We performed a search of the annotated proteins in RefSeq to find small proteins (less than 60 amino acids in length) that are encoded directly by small ORFs and not processed from a larger protein. In the mouse, only 26 proteins were found that met these criteria. Of these 26, only 16 (including three ribosomal proteins) were confirmed to produce proteins in vivo.

*TUG1* represents one of a limited number of ncRNAs that have been shown to have a function in vivo. However, how *TUG1* is acting as a ncRNA is unclear. *TUG1* is a spliced and polyadenylated polymerase-II transcript that does not appear to be processed into a miRNA. As a potential ncRNA, *TUG1* may function globally to alter chromatin configurations to aid Crx and Nrl in the activation of photoreceptor-specific genes in the retina. It is intriguing that, on the 3′ side of *TUG1*, there is an overlapping locus, TI-227, which a study of metastatic melanoma cells found to be altered in its level of expression. A recent study of a ncRNA in yeast found that a ncRNA regulated the transcription of an adjacent locus through transcription interference [43]. Further studies to understand the mechanism of *TUG1*

action will be required in order to explore this and other mechanisms that are employed by the thousands of putative ncRNAs that are expressed in higher vertebrates.

## Experimental Procedures

### Animals
Timed pregnant Sprague Dawley rats were purchased from Taconic farms. Animals were treated as per the guidelines set by the Institutional Animal Care and Use committee at Harvard University.

### Northern-Blot Analyses of m Transcripts
Standard Northern-blot practices (Ambion) were performed for analysis of *mTUG1* transcripts in the developing brain and retina. Ten micrograms of total RNA was loaded into each well of a 1% formaldehyde gel. After transfer to a nylon membrane, blots were probed for *mTUG1* transcripts. Probes used for analysis included two double-stranded-DNA probes labeled with [$\alpha$-$^{32}$P]dCTP and the Rediprime kit (Amersham). These probes were: *mTUG1* full length (nt 1–6, 721 of *TUG1c*) and *mTUG1* partial (nt 203–1688 of *mTUG1c*) (see Figure 2F). For strand specificity, an RNA probe was synthesized by in vitro transcription with the MaxiScript Kit (Ambion). This probe was synthesized in the – direction, opposite to that of *TUG1a–c* from nt 1020–4628 (*mTUG1c*) (see Figure 2F).

### mTUG1 cDNA Cloning
Full-length *mTUG1c* was cloned by ligation of three PCR products and two EST clones. These *mTUG1* fragments include: EST AI842058, EST BE981391, TUG partial (nt 203–1688), *TUG1* beg (nt 1020–4628), and TUG end (nt 3040–6585). PCR products were amplified from E16 and P0 mouse total-brain cDNA or from a bacterial artificial chromosome (BAC) containing the *TUG1* locus (Genbank ID AL691413). Sequencing of multiple clones of full-length cloned *mTUG1c* confirmed the predicted genomic sequence and predicted Unigene sequence (Mm.359796).

### RNase Protection Assays
RNase protection assays were performed with the RPA III kit from Ambion. Ten micrograms of total RNA were used for target RNA. The utilized mouse *TUG1* probes covered the following sequences: *TUG1* RPA probe, nt 3040–3461 of *TUG1c*, transcribed from the + strand in the – direction, as indicated. For $\beta$-actin, a control probe that recognizes a protected fragment of 188 nt was also used (Ambion). RNA probes were synthesized via standard in vitro-transcription protocols with [$\alpha$-$^{32}$P] (Ambion). Samples were electrophoresed on a 6% TBE-Urea gel. After electrophoresis, gels were dried and exposed to Hyperfilm (Amersham).

### In Vivo Electroporations
In vivo electroporations were performed at P0 as described [20], and retinas were harvested at P3, P4, P21, or P25. The *TUG1a* RNAi target sequence was 5′-GGGCCAGAGACAAATGGTTTCC-3′; the *TUG1b* RNAi target sequence was 5′-GGGTTACTCAGGAAC CAAAAC-3′. RNAi constructs were cloned into the pBSU6 vector [21] and coelectroporated with membrane bound GFP driven by a ubiquitous CAG promoter [44]. One microliter of DNA was injected into the subretinal space of Sprague Dawley rats. Each DNA was electroporated at 1 mg/ml.

For microarray experiments, two independent experiments were performed. For each of these experiments, three P0 retinas per variable were coelectroporated with CAG-mGFP and either *TUG1* RNAi or an empty U6 vector. At P3, retinas were harvested, dissociated with papain, and sorted by FACs to collect GFP-positive cells. RNA was prepared via the Trizol method (GibcoBRL), and cDNA was synthesized and amplified via the SMART amplification protocol (Clontech). Amplified cDNA was directly labeled with Cy3 or Cy5 and hybridized to a cDNA microarray as described [45]. Slides were scanned in an Axon GenePix 4000B Scanner (Axon Instruments), and data were extracted from the resulting images with the GenePix software package (Axon Instruments).

For immunocytochemistry, at least three independent P0 retinas per variable were electroporated in vivo as described above and harvested at P25. Retinas were fixed in 4% paraformaldehyde for 30 min, cryoprotected with 30% sucrose in PBS overnight at 4°C and then 1–2 hr in 1:1 30% sucrose:OCT, and embedded in OCT (TissueTek). Retinas were then sectioned (20 $\mu$m) and stained with antibodies as described below.

### In Situ Hybridizations and Immunocytochemistry
In situ hybridization and immunofluorescent staining were performed on cryosections (20 $\mu$m). Section in situ hybridizations were performed as described [46]. Probes used for in situ analysis were *mTUG1* (nt 436–1072 of and nt 3040–6585), each synthesized in the direction opposite to *TUG1a–c*. These probes should recognize *mTUG1a–c*. Images were taken on a Nikon Eclipse E1000 microscope with a Leica DC200 digital camera.

For immunofluorescent staining, retinal cryosections (20 $\mu$m) or dissociated retinal cells were blocked with 2% goat serum/2% donkey serum/0.1% Triton X-100 for 1 hr at room temperature. Slides were incubated in primary antibody overnight at 4°C. The primary antibodies used were: anti-rhodopsin, Rho4D2 (mouse monoclonal, 1:250 [47]), anti-Chx10 (rabbit polyclonal, 1:500 [A. Chen and C.L.C., unpublished data]), anti-Pax6 (mouse monoclonal, 1:500, University of Iowa Developmental Studies Hybridoma Bank, deposited by Atsushi Kawakami), anti-cone opsins (rabbit polyclonals, 1:1000, gift of J. Nathans), anti-PNA (rhodamine-conjugated, 1:1000, Vector Labs), anti-glutamine synthetase (mouse monoclonal, 1:200, PharMingen), and anti-GFP (mouse monoclonal, rabbit polyclonal, both 1:500, Molecular Probes). After several washes (PBS, 0.1% Triton X-100), slides were incubated in goat anti-mouse or goat anti-rabbit Cy3 or Cy5 (Jackson Immunoresearch Laboratory 1:200) for 1–2 hr at room temperature. Cells were counterstained with DAPI and washed several times. Dissociated samples were visualized and quantified on a Nikon Eclipse E1000 microscope. Immunofluorescent stainings of retinal cryosections were photographed on an Axioplan 2 microscope with a LSM 510 Meta module for confocal imaging.

### TUNEL Assays
P0 rat retinas were electroporated as described above. At P4, retinas were harvested and sectioned as described above. TUNEL assays were performed on retinal cryosections (20 $\mu$m) as per manufacturer's instructions (In Situ Cell Death Detection Kit, TMR red, Roche).

## References

1. Sturman, J.A. (1993). Taurine in development. Physiol. Rev. *73*, 119–147.
2. Hageman, G.S., and Schmidt, S.Y. (1987). Taurine-deficient pigmented and albino rats: Early retinal abnormalities and differential rates of photoreceptor degeneration. Prog. Clin. Biol. Res. *247*, 497–515.
3. Imaki, H., Moretz, R., Wisniewski, H., Neuringer, M., and Sturman, J. (1987). Retinal degeneration in 3-month-old rhesus monkey infants fed a taurine-free human infant formula. J. Neurosci. Res. *18*, 602–614.
4. Sturman, J.A., Gargano, A.D., Messing, J.M., and Imaki, H. (1986). Feline maternal taurine deficiency: Effect on mother and offspring. J. Nutr. *116*, 655–667.

5. Altshuler, D., Lo Turco, J.J., Rush, J., and Cepko, C. (1993). Taurine promotes the differentiation of a vertebrate retinal cell type in vitro. Development *119*, 1317–1328.

6. Young, T.L., and Cepko, C.L. (2004). A role for ligand-gated ion channels in rod photoreceptor development. Neuron *41*, 867–879.

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

8. Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., and Swaroop, A. (2001). Nrl is required for rod photoreceptor development. Nat. Genet. *29*, 447–452.

9. Furukawa, T., Morrow, E.M., and Cepko, C.L. (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. Cell *91*, 531–541.

10. Tanaka, S. (2003). Comparative aspects of intracellular proteolytic processing of peptide hormone precursors: Studies of proopiomelanocortin processing. Zoolog. Sci. *20*, 1183–1198.

11. Mattick, J.S. (2003). Challenging the dogma: The hidden layer of non-protein-coding RNAs in complex organisms. Bioessays *25*, 930–939.

12. Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., et al. (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature *420*, 563–573.

13. Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. (2002). Xist RNA and the mechanism of X chromosome inactivation. Annu. Rev. Genet. *36*, 233–278.

14. Lee, J.T., Davidow, L.S., and Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. Nat. Genet. *21*, 400–404.

15. Hurst, L.D., and Smith, N.G. (1999). Molecular evolutionary evidence that H19 mRNA is functional. Trends Genet. *15*, 134–135.

16. Ishiguro, T., Nakajima, M., Naito, M., Muto, T., and Tsuruo, T. (1996). Identification of genes differentially expressed in B16 murine melanoma sublines with different metastatic potentials. Cancer Res. *56*, 875–879.

17. Ishiguro, T., Nagawa, H., Naito, M., Tsuruo, T., Nakajima, M., and Muto, T. (2000). Analysis of novel metastasis-associated gene TI-227. Jpn. J. Cancer Res. *91*, 390–394.

18. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

19. Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

20. Matsuda, T., and Cepko, C.L. (2004). Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc. Natl. Acad. Sci. USA *101*, 16–22.

21. Sui, G., Soohoo, C., Affar, E.B., Gay, F., Shi, Y., and Forrester, W.C. (2002). A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proc. Natl. Acad. Sci. USA *99*, 5515–5520.

22. Choi, D.K., Ito, T., Mitsui, Y., and Sakaki, Y. (1998). Fluorescent differential display analysis of gene expression in apoptotic neuroblastoma cells. Gene *223*, 21–31.

23. Choi, D.K., Ito, T., Tsukahara, F., Hirai, M., and Sakaki, Y. (1999). Developmentally-regulated expression of mNapor encoding an apoptosis-induced ELAV-type RNA binding protein. Gene *237*, 135–142.

24. Trioulier, Y., Torch, S., Blot, B., Cristina, N., Chatellard-Causse, C., Verna, J.M., and Sadoul, R. (2004). Alix, a protein regulating endosomal trafficking, is involved in neuronal death. J. Biol. Chem. *279*, 2046–2052.

25. Chatellard-Causse, C., Blot, B., Cristina, N., Torch, S., Missotten, M., and Sadoul, R. (2002). Alix (ALG-2-interacting protein X), a protein involved in apoptosis, binds to endophilins and induces cytoplasmic vacuolization. J. Biol. Chem. *277*, 29108–29115.

26. Cregan, S.P., Dawson, V.L., and Slack, R.S. (2004). Role of AIF in caspase-dependent and caspase-independent cell death. Oncogene *23*, 2785–2796.

27. Haas, H.L., and Hosli, L. (1973). The depression of brain stem neurones by taurine and its interaction with strychnine and bicuculline. Brain Res. *52*, 399–402.

28. Horikoshi, T., Asanuma, A., Yanagisawa, K., Anzai, K., and Goto, S. (1988). Taurine and beta-alanine act on both GABA and glycine receptors in Xenopus oocyte injected with mouse brain messenger RNA. Brain Res. *464*, 97–105.

29. Lewis, C.A., Ahmed, Z., and Faber, D.S. (1991). A characterization of glycinergic receptors present in cultured rat medullary neurons. J. Neurophysiol. *66*, 1291–1303.

30. Schmieden, V., Kuhse, J., and Betz, H. (1992). Agonist pharmacology of neonatal and adult glycine receptor alpha subunits: Identification of amino acid residues involved in taurine activation. EMBO J. *11*, 2025–2032.

31. Wu, Z.Y., and Xu, T.L. (2003). Taurine-evoked chloride current and its potentiation by intracellular Ca(2+) in immature rat hippocampal CA1 neurons. Amino Acids *24*, 155–161.

32. Rivera, C., Voipio, J., Payne, J.A., Ruusuvuori, E., Lahtinen, H., Lamsa, K., Pirvola, U., Saarma, M., and Kaila, K. (1999). The K+/Cl- co-transporter KCC2 renders GABA hyperpolarizing during neuronal maturation. Nature *397*, 251–255.

33. Flint, A.C., Liu, X., and Kriegstein, A.R. (1998). Nonsynaptic glycine receptor activation during early neocortical development. Neuron *20*, 43–53.

34. LoTurco, J.J., Owens, D.F., Heath, M.J., Davis, M.B., and Kriegstein, A.R. (1995). GABA and glutamate depolarize cortical progenitor cells and inhibit DNA synthesis. Neuron *15*, 1287–1298.

35. Sturman, J.A., Moretz, R.C., French, J.H., and Wisniewski, H.M. (1985). Taurine deficiency in the developing cat: Persistence of the cerebellar external granule cell layer. Prog. Clin. Biol. Res. *179*, 43–52.

36. Sturman, J.A., Palackal, T., Imaki, H., Moretz, R.C., French, J., and Wisniewski, H.M. (1987). Nutritional taurine deficiency and feline pregnancy and outcome. Adv. Exp. Med. Biol. *217*, 113–124.

37. Carter-Dawson, L.D., LaVail, M.M., and Sidman, R.L. (1978). Differential effect of the rd mutation on rods and cones in the mouse retina. Invest. Ophthalmol. Vis. Sci. *17*, 489–498.

38. Nishida, A., Furukawa, A., Koike, C., Tano, Y., Aizawa, S., Matsuo, I., and Furukawa, T. (2003). Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. Nat. Neurosci. *6*, 1255–1263.

39. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

40. Zhuo, D., Zhao, W.D., Wright, F.A., Yang, H.Y., Wang, J.P., Sears, R., Baer, T., Kwon, D.H., Gordon, D., Gibbs, S., et al. (2001). Assembly, annotation, and integration of UNIGENE clusters into the human genome draft. Genome Res. *11*, 904–918.

41. Kozak, M. (1996). Interpreting cDNA sequences: Some insights from studies on translation. Mamm. Genome *7*, 563–574.

42. Chan, Y.L., Olvera, J., and Wool, I.G. (1995). The primary structures of rat ribosomal proteins L4 and L41. Biochem. Biophys. Res. Commun. *214*, 810–818.

43. Martens, J.A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *429*, 571–574.

44. Niwa, H., Yamamura, K., and Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene *108*, 193–199.

45. Livesey, F.J., Furukawa, T., Steffen, M.A., Church, G.M., and Cepko, C.L. (2000). Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. Curr. Biol. *10*, 301–310.

46. Chen, C.M., and Cepko, C.L. (2002). The chicken RaxL gene plays a role in the initiation of photoreceptor differentiation. Development *129*, 5363–5375.

47. Molday, R.S., and MacKenzie, D. (1983). Monoclonal antibodies to rhodopsin: Characterization, cross-reactivity, and application as structural probes. Biochemistry *22*, 653–660.