Structural Organization and Expression Pattern of the Canine *RPGRIP1* Isoforms in Retinal Tissue

Tatyana Kuznetsova,¹ Barbara Zangerl,¹ Orly Goldstein,² Gregory M. Acland,² and Gustavo D. Aguirre¹

PURPOSE. To examine the structure and expression of *RPGRIP1* in dog retina.

METHODS. Determination of the structural analysis and expression pattern of canine *RPGRIP1* (*crRPGRIP1*) was based on cDNA amplification. Absolute quantification of the expression level of *crRPGRIP1* splice variants was determined by qRT-PCR. Regulatory structures were examined by computational analysis of comparative genomics.

RESULTS. *crRPGRIP1* encompasses 25 exons that harbor a 3627-bp open reading frame (ORF) encoding a 1209-amino-acid (aa)-predicted protein. In addition to the main transcript, five full-length and several partial *crRPGRIP1* isoforms were identified revealing four alternative 3'-terminal exons—24, 19a, 19c, and 19d—three of which could potentially produce C-terminally truncated proteins that lack the RPGR-interacting domain. A complex organization of the 5'-UTR for the *crRPGRIP1* splice products have been described, with a common promoter driving multiple isoforms, including four full-length transcripts using the 3'-terminal exon 24. In addition, a potential alternative internal promoter was revealed to initiate at least two *crRPGRIP1* splice variants sharing the same 3'-terminal exon 19c. Transcription initiation sites were highly supported by conserved arrangements of *cis*-elements predicted in a bioinformatic analysis of orthologous *RPGRIP1* promoter regions.

CONCLUSIONS. The use of alternative transcription start and termination sites results in substantial heterogeneity of *crRPGRIP1* transcripts, many of which are likely to have tissue-specific expression. The identified exon-intron structure of *crRPGRIP1* isoforms provides a basis for evaluating the gene defects underlying inherited retinal disorders in dogs. (*Invest Ophthalmol Vis Sci.* 2011; 52:2989–2998) DOI:10.1167/iov.10-6094

The *RPGRIP1* gene encodes retinitis pigmentosa GTPase-interacting protein 1, and mutations in the human gene are associated with Leber congenital amaurosis (LCA),^{1–3} juvenile

retinitis pigmentosa,⁴ a late-onset cone-rod dystrophy,⁵ and cone-rod dystrophy type 1 in dogs.^{6,7} Human *RPGRIP1* consists of 25 exons, of which 24 code for a 1259-amino-acid protein.^{1,3} Exons 6 to 13, 14 to 16, and 18 to 24 encode, respectively, the α -helical coil-coiled protein interaction motif of members of the structural maintenance of chromosomes (SMC) superfamily, two protein kinase C conserved region 2 motifs (C2), and conserved RPGR-interacting domain (RID).^{8–10}

The biological functions of *RPGRIP1* are complex. In the eye, it is expressed in amacrine neurons^{8,11} and photoreceptors,^{9,12} and in numerous other tissues, albeit at greatly reduced levels.^{9,12} Moreover, the existence of multiple isoforms, with species-specific subcellular localization patterns (e.g., connecting cilium,^{11,13} photoreceptor inner¹⁴ and outer^{8,11} segments, and basal bodies of cells with primary cilia¹⁵), suggests that different isoforms perform cell-specific functions. *RPGRIP1* is required not only for disc morphogenesis of the outer segments (OS),¹⁶ but also for the formation of the OS itself, particularly in rods.¹⁷ A general role of *RPGRIP1* as a scaffold protein has been suggested,¹³ and it interacts directly or indirectly with *RPGR*,^{9,16,18} *NPHP4*,¹⁰ and *RanBP2*.⁸ It has been shown that *RPGRIP1* can be proteolytically processed, rendering its N-terminal domain competent for nuclear localization,¹⁹ suggesting that it may be involved in regulating gene expression.

Although a mutation in *RPGRIP1* is causally associated with canine cone-rod dystrophy,^{6,7} a potential large-animal model for gene-based therapies,²⁰ little is known about the canine gene structure, organization, and expression, and the molecular basis of the disease. To assess the structure/function relationship of the *RPGRIP1* isoforms, we characterized the full-length transcript of canine *RPGRIP1* (*crRPGRIP1*) and its several alternatively spliced isoforms and evaluated the 5'- and 3'-UTRs of *RPGRIP1* transcripts. Our results identified a novel complex 5' and 3' splicing pattern and further described the complete structure of six *crRPGRIP1* alternatively spliced variants driven by two different promoters.

MATERIALS AND METHODS

Tissue Sources and cDNA Synthesis

The research was conducted in full compliance with the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research. Archived tissue samples were obtained from the Research Repository at the RDS Facility, University of Pennsylvania. Total RNA was isolated from canine tissues by using an extraction reagent (TRIzol; Invitrogen, Carlsbad, CA) and single chloroform extraction. The quality of the RNAs was evaluated by UV spectroscopy and denaturing formaldehyde-agarose gel. First-strand cDNA was synthesized in 20- μ L reactions by using random hexamers primers and reverse transcriptase (SmartScribe; Clontech, Mountain View, CA) according to the manufacturer's recommendations.

From the ¹Section of Ophthalmology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; and the ²James A. Baker Institute for Animal Health, Cornell University, Ithaca, New York.

Supported by Morris Animal Foundation; National Eye Institute, National Institutes of Health Grants EY-06855, EY-13132, EY-17549, and P30 EY-001583; Foundation Fighting Blindness center grant; the Van Sloun Fund for Canine Genetic Research; Hope for Vision; The ONCE International Prize for R&D in Biomedicine and New Technologies for the Blind.

Submitted for publication June 21, 2010; revised October 26 and November 30, 2010; accepted December 1, 2010.

Disclosure: T. Kuznetsova, None; B. Zangerl, None; O. Goldstein, None; G.M. Acland, None; G.D. Aguirre, None

Corresponding author: Gustavo D. Aguirre, School of Veterinary Medicine, University of Pennsylvania, 3900 Delancey Street, Philadelphia, PA 19104; gda@vet.upenn.edu.

Amplification of Long PCR Fragments and Sequencing

Amplification of *cRPGRIP1* cDNA fragments was performed with gene-specific primers: 2/3EX F (or 2/5EX F) and 22EX R (for primer sequences, see Supplementary Table S1, <http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10.6094/-DCSupplemental>). PCR was performed with a commercial system (Expand Long Template PCR System; Roche, Indianapolis, IN), and the products were visualized on 1.5% agarose gels. PCR conditions were 4 minutes of initial denaturation at 95°C followed by 10 cycles of 95°C for 20 seconds, 60°C for 30 seconds (annealing), 68°C for 2.5 minutes (extension); the next 25 cycles were 95°C for 20 seconds, 60°C for 30 seconds, and 68°C for 2.5 minutes, with a 20-second cycle elongation for each successive cycle.

Long-range PCR performed with 2/3EX F and 22EX R primers on retinal mRNA produced few PCR fragments; an approximately 3100-bp PCR fragment was isolated and purified from the agarose gel with an extraction kit (NucleoTrap; Clontech) and subjected to reamplification, resulting in a single band. Then, using this PCR product as a template, we generated and sequenced overlapping PCR products. The sequences were aligned and analyzed (Lasergene software, ver. 7.2.1 DNASTar, Madison, WI, and Vector NTI, ver. 10.3.0; Invitrogen, Carlsbad, CA).

PCR performed with 2/5EX F and 22EX R produced two clearly visible PCR fragments whereas the lower molecular weight of the two products was composed of two bands (1357/1390 bp) indistinguishable on a 1.5% agarose gel because of their similar size. Both PCR fragments were isolated and purified from the gel, processed, and sequenced, as described.

5' and 3'-RACE

5'-RACE was performed with the RNA-ligase-mediated RACE (RLM-RACE) system (Ambion, Austin, TX), according to the manufacturer's recommendations and using gene-specific primers listed in Supplementary Table S1 (<http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10.6094/-DCSupplemental>). The 5' end of the *RPGRIP1* transcripts was obtained in two steps using the gene-specific outer (10/11EX R, 13/19EX R, or 13/19cEX R) and inner (1A/2EX R, 2/7EX R, 1/10EX R, 2/4EX R, 2/5EX R, EX3 R, or 12EX R) primers. Cycling conditions were 4 minutes of initial denaturation at 95°C followed by 35 cycles of 95°C for 20 seconds, 54°C for 30 seconds (annealing), and 68°C for 2.5 minutes (extension). Amplified products, usually a single band, were

analyzed on a 2% agarose gel, cut from the gel, extracted (NucleoTrap Gel Extraction Kit; Clontech), and directly sequenced.

For 3'-RACE, we used the adapter-specific primer (5'-CGATAGGGCAAGCAGTGGTATCAACGCAGAGCAC(T₁₇)G) for first-strand cDNA synthesis and reverse transcriptase (SmartScribe; Clontech). The 3' end of *cRPGRIP1* transcripts were amplified with adapter specific outer primer UP1, and a gene-specific primer located in coding exon 19 (19EX F), across the exon 21-exon 22 junction (21/22EX F), or in exon 12 (12EX F). Cycling conditions were 4 minutes of initial denaturation at 95°C followed by 35 cycles of 95°C for 20 seconds and 68°C for 3 minutes (annealing and extension). 3'-RACE PCR products were analyzed on a 1.5% agarose gel, extracted (NucleoTrap Gel Extraction Kit; Clontech) and cloned with a kit (TOPO TA; Invitrogen). The inserts were sequenced with standard M13 forward and reverse primers.

Expression Profile Analysis by RT-PCR

Analysis was performed with canine retinal cDNA (3, 7.4, and 16 weeks, and 6 months); brain (temporal [TL], frontal [FL], and occipital [OL] lobes), lung, testis, and heart cDNA (6 months in all cases), and was used to estimate expression level of *cRPGRIP1*. Gene-specific primers are shown in Table 1. Cycling conditions were 4 minutes of initial denaturation at 95°C, followed by 35 cycles of 95°C for 20 seconds, 60°C for 30 seconds (annealing), and 68°C for 20 seconds (extension). PCR products were analyzed on a 6% polyacrylamide gel followed by ethidium bromide staining.

Absolute Quantification of *RPGRIP1* Transcripts in Retina

Real-time PCR was performed in a total volume of 25 μ L in 96-well microwell plates on a real-time PCR system (model 7500; Applied Biosystems, Inc. [ABI], Foster City, CA). The amplification data were analyzed with the system software (7500 ver. 2.0.1; Applied Biosystems, Inc., [ABI], Foster City, CA). Specific RT-PCR product was used for construction of an absolute standard curve for individual amplicons representing the characterized reference sequence, and isoforms 2-5a, 2-5b, 2-5c, 19c-1, and 19c-2. The number of copies of a template was calculated based on the assumption that the average molecular weight of a base pair (bp) is 650. Quantity of an amplicon (ng/ μ L) was assessed by UV spectroscopy and inversely to the number of moles of template present in 1 μ L of material. Using Avogadro's number,

TABLE 1. Primers Sequences

Primer	Primer Sequence (5'–3')	Amplicon Name (Exons and Size, bp)	Transcript	Application
13L/20EX, F	GGCCTGCCACTACCG GA TAAAGAAT	2-5c (13L-20-21; 151 bp)	2-5c	RT-PCR, real-time PCR
21EX, R	CCAGCGAGATCACTTCAATGCACA			
13/20EX, F	GCATCCGTCAGCTGGAA GA TAAAGAAT	2-5b (13-20-21; 153 bp)	2-5b	RT-PCR, real-time PCR
21EX, R	CCAGCGAGATCACTTCAATGCACA			
5/6EX, F	CAGAGCTGC GA CTTCCATTAAGGATAA	2-5a (5-6-7-10; 150 bp)	2-5a	RT-PCR, real-time PCR
7/10EX, R	GGCCAGCAT CT TCTGGAGCAGA			
14/15EX, F	ACCACACTTACT CG AGCTGGTGAGAA	14-16EX (14-15-16; 232 bp)	Reference sequence	RT-PCR, real-time PCR
16EX, R	GGAGGCCGTGG CA TCTGATGA			
1/2EX, F	CAGTGTGCAAA CG TAAAGAATATGAAAGCTCAAC	1-3EX (1-2-3; 178 bp)	Reference sequence	RT-PCR, real-time PCR
3EX, R	GCAGCCGTCAGCCGGAGCAG			
1aEX, F	ACAGGACCAGAGGACACCAGAACAG	1A-2EX (1A-1-2; 289 bp)		RT-PCR
2EX, R	CTCCTTCACCAACATGTGCTCTTCAC			
19EX, F	CAGGTGAACACTCACTGAATGGAAGTTCT	19-21EX (19-20-21; 273 bp)		RT-PCR
21EX, R	CAGCGAGATCACTTCAATGCACA			
12/13EX, F	TTACCA CG AGGAAGTGAAGCAGTGAA	19c-1 (12-13-19c; 140 bp)	19c-1	RT-PCR, real-time PCR
13/19cEX, R	TGGTACCAGTTGT CT TCCAGCTGAC			
12EX, F	GGCCCTGCAGCGAGCCATCAA	19c-2 (12-13-19; 168 bp)	19c-2	RT-PCR, real-time PCR
13/19EX, R	CAGTGTAGTTACCT CT TCAGCTGAC			

Exon-exon boundary in a primer is shown in the boxed region. Note that the 3'-end nucleotide for primer 12/13EX, F (A, bold) was replaced from original sequence that contains -T in this position.

6.022×10^{23} molecules/mole, the number of molecules of the template in 1 μ L was calculated. The dynamic range of the calibration curves was at the start between 10^2 and 10^7 molecules. All PCRs were performed in microwells using cDNA generated from 50 ng of DNase-treated retinal RNA (6 month old normal dog), 12.5 μ L SYBR green PCR Master Mix, 0.15 μ M both forward and reverse primer and DEPC treated water to a 25 μ L volume. Analysis was performed on two plates; the first plate contained the reference, 2-5a, 2-5b, 2-5c and the second the reference, 19c-1, and 19c-2.

Promoter Sequences

An orthologous gene approach was used to identify promoter sequences for *RPGRIP1*. Based on structural similarity and high homology between orthologous promoter regions, the canine sequence (1000 bp upstream to 100 bp downstream of the *cRPGRIP1* of exon 1A or coding exon 10) was aligned with bovine, human, and murine sequences using UCSC Genome Browser (BLAT; <http://genome.ucsc.edu/> provided in the public domain by UCSC Genome Bioinformatics, University of California at Santa Cruz, Santa Cruz, CA). Also NM_181034 and AF265669 (bovine), NM_023870 (murine), NM_020366 and BX646036 (human) orthologous sequences were used for comparative analysis. Sequence distances were estimated (Lasergene software ver. 7.2.1; DNASTar), and a computational analysis of promoter sequences was performed (GEMS Launcher software; FrameWorker algorithm, which identifies common elements in DNA sequences from different species, and MatInspector, which searches for DNA-binding sites; Genomatix GmbH, Munich, Germany; <http://www.genomatix.de/>) with analysis restricted to transcription factors relevant to the eye.

www.genomatix.de/) with analysis restricted to transcription factors relevant to the eye.

RESULTS

Characterization of *cRPGRIP1* Gene Structure

The complete sequence of *cRPGRIP1* has not been characterized fully. To date, most ESTs in the nucleotide databases are derived from cDNA fragments from testis and represent either the 3' end (exons 20–24 and 3'-UTR) or the middle part of the gene (splice variants comprising exons 13–17). Two ESTs (DT536462 and DT540981) represent retinal *RPGRIP1* cDNAs obtained from a normalized canine retinal cDNA library²¹ and include exons 21 to 23 and a partial exon 24 sequence. In addition, 62 nucleotides of coding sequence of *cRPGRIP1* (fragments of exons 2 and 3) have been published.⁶ This information was insufficient to draw any conclusion as to the structure of *cRPGRIP1* for which, based on the data for orthologous genes, multiple splice isoforms are expected.

As a first step, human and bovine cDNA sequences were aligned with the canine genomic sequence to determine the most likely positions of the exon-intron boundaries. Areas of homology were used for primer design, to amplify overlapping fragments spanning the predicted translated region. This approach identified cDNA fragments representing multiple alternatively spliced gene products, especially the region between

TABLE 2. Characterization of Exon Structure of Canine *RPGRIP1*

Exon	Length (bp)	Chromosome Localization	Regulatory Features	Intron Size (bp)
1A	95/68	<i>crb15</i> (+): 21,323,665–21,323,759	TSS2: 21,323,665 TSS1: 21,323,692	6,726
1	120	<i>chr15</i> : 21,330,486–21,330,605		8,086
2	133	<i>chr15</i> : 21,338,692–21,338,824		3,900
3	224	<i>chr15</i> : 21,342,725–21,342,948		938
4	91	<i>chr15</i> : 21,343,887–21,343,977		967
5	114	<i>chr15</i> : 21,344,945–21,345,058		1,890
6	106	<i>chr15</i> : 21,346,949–21,347,054		3,656
7	24	<i>chr15</i> : 21,350,711–21,350,734		1,074
8	147	<i>chr15</i> : 21,351,809–21,351,955		475
9	74	<i>chr15</i> : 21,352,431–21,352,504		3,132
10	146	<i>chr15</i> : 21,355,637–21,355,782	TSS3: 21,355,637	3,047
11	140	<i>chr15</i> : 21,358,830–21,358,969		133
12	144	<i>chr15</i> : 21,359,103–21,359,246		457
13	118	<i>chr15</i> : 21,359,704–21,359,821		2,232
14	453	<i>chr15</i> : 21,362,054–21,362,506		160
15	152	<i>chr15</i> : 21,362,667–21,362,818		440
16	343	<i>chr15</i> : 21,363,259–21,363,601		1,202
17	161	<i>chr15</i> : 21,364,804–21,364,964		329
18	204	<i>chr15</i> : 21,365,294–21,365,497		1,818
19	139	<i>chr15</i> : 21,367,316–21,367,454		10,712
20	98	<i>chr15</i> : 21,378,167–21,378,264		4,905
21	193	<i>chr15</i> : 21,383,170–21,383,362		5,765
22	85	<i>chr15</i> : 21,389,128–21,389,212		2,666
23	125	<i>chr15</i> : 21,391,879–21,392,003		2,126
24	194	<i>chr15</i> : 21,394,130–21,394,323	Stop codon: 21,394,324–21,394,326	
13L	151	<i>chr15</i> : 21,359,704–21,359,854		
19a	251	<i>chr15</i> : 21,370,770–21,371,020	Stop-codon: 21,371,021–21,371,023	
19b	39	<i>chr15</i> : 21,372,473–21,372,511		
19c	89	<i>chr15</i> : 21,374,879–21,374,967	Stop-codon: 21,374,968–21,374,970	
19d	62	<i>chr15</i> : 21,376,538–21,376,599	Stop-codon: 21,376,600–21,376,602	

The reference transcript has 25 exons (24 coding exons). The transcription start sites (TSS1, TSS2, and TSS3) and predicted stop codons are noted. Sequences for alternative exons are shown in bold.

exons 2 and 19, and the data were assembled to create a preliminary gene model used to design primers for the subsequent 3'- and 5'-RACE experiments. Based on the high degree of homology of *RPGRIP1* sequences with orthologous genes, we predicted that the full-length *cRPGRIP1* transcript encompasses 25 exons, of which 24 are coding (Table 2, Fig. 1A).

To confirm the prediction, we designed primers across the exon 2-exon 3 junction and a reverse primer in exon 22 for long-range PCR on retinal mRNA and obtained a 3105-bp product. A gene-specific primer was placed in exon 3, and the 5' end was obtained by RLM 5'-RACE, which selectively amplifies only mRNA with a 5' cap. By placing a gene-specific primer in exon 19, as a gene-specific primer, we obtained exons 19 to 24 and 3'-UTR by 3'-RACE. The results indicated that *cRPGRIP1* has an open reading frame ORF of 3627 bp that encodes a predicted protein product of 1209 amino acids (aa) (GenBank accession number HM021768; <http://www.ncbi.nlm.nih.gov/Genbank>; provided in the public domain by the National Cen-

ter for Biotechnology Information, Bethesda, MD). We refer to this as the reference sequence.

To characterize the 5' region, we used a forward primer in noncoding exon 1A combined with a reverse primer across the exon 10-exon 11 junction and performed PCR to estimate a sequence variability in this region. We found that noncoding exon 1A is common to a number of 5'-splice gene products (Fig. 1B), but for most of them, the complete structure is still to be elucidated. Results of the RLM 5'-RACE, however, showed that both the reference sequence and its splicing isoforms, share the same transcription start site (designated as TSS1; Table 2), and seem to be regulated by the same promoter (P1). A putative tata-like element (–tattaat–) was found at position –43 to –37 from TSS1. In addition to TSS1, a minor alternative TSS (TSS2; Table 2), located 27 bp upstream from TSS1, was identified in the 2-4 isoform.

The 3'-RACE experiment, undertaken with a gene-specific primer located in exon 19, showed that *cRPGRIP1* transcripts

FIGURE 1. (A) Exon structure of the canine *RPGRIP1* reference sequence. The gene consists of 24 coding exons (numbers in boxes). The TSS is shown by an arrow at +1 and the predicted translation start site by an arrow, with Met indicated for the transcript. The number of the last full amino acid (aa) at the end of each exon appears below each exon box. Exons 1 and 2, part of exon 3, and 18 to 24 encode, respectively, the ND¹⁹ and the conserved RID.^{9,10} Direct analysis of canine *RPGRIP1* sequence with the protein BLAST tool identified two structurally different regions that were described for orthologous sequences: the SMC/CC-domain encoded by exons 6 to 13 and C2-domain encoded by exon 16. (B) Noncoding exon 1A is common to new *cRPGRIP1* isoforms (exons 1A-10 are represented). The transcription (arrow shown at +1 on all isoforms) and predicted translation (arrow with Met indicated for each isoform) initiation sites are noted. (C, D) Alternative 3'-terminal exons and PASs of *cRPGRIP1*. Location of alternative 3'-terminal exons (19a, 19b, 19c, 24) and PAS1 to 4 identified in this study. Ethidium bromide-stained agarose gel of 3'-RACE showing the 3' isoforms. *Heteroduplex DNA. Dashed arrows: location of PCR products barely visible on the figure but visible on the gel.

arise from a common pre-mRNA by alternative splicing of four 3'-terminal exons. Sequencing of the 3'-RACE products revealed four different polyadenylation sites in the sequence that give rise to transcripts with four different 3'-UTRs after the 3'-terminal exon (exons 19a, 19c, 19d, or 24; Figs. 1C, 1D, Table 2). However, the 3'-UTR sequence after exon 19c can be potentially incomplete, since a stretch of oligo-A (A_{15}) is also present in the DNA sequence, suggestive of putative mispriming of polyT-primer and a likely cause of the artifactual PAS2. Polyadenylation signal motifs are in close proximity to all four poly(A) sites (PAS1–PAS4). A consensus polyadenylation signal –AATAAA– was present at position –25 to –20 from PAS1. Putative polyadenylation signals, –AATTAA– at position –29 to –23 from the PAS2 and at position –24 to –19 from PAS3, and –AACAAA– at position –14 to –9 from the PAS4, were also identified. Notably, transcripts terminating in alternative exon 19a, 19c, or 19d will lead to the premature truncation of *RPGRIP1* and the deletion of its RID domain, likely affecting interactions with *RPGR*.

Structural Characterization of *cRPGRIP1* Isoforms 2-5a, 2-5b, 2-5c, 19c-1, and 19c-2

Full-length cDNA was characterized for the retina-expressed, alternatively spliced *cRPGRIP1* isoforms 2-5a, 2-5b, and 2-5c (GenBank accession numbers HM021769, HM021770, and HM021771, respectively) that create downstream reading frames shorter than the reference sequence (Fig. 2). The three isoforms were identified by RT-PCR with a forward primer spanning the exon 2–exon 5 junction, and a reverse primer located in exon 22, combined with 5'- and 3'-RACE, by using the gene-specific primers 2/5EX R and 21/22EX F, respectively. Analysis of the 2-5a isoform structure indicates that it splices out exons 3, 4, 8 to 9, and 14 to 16, and encodes a predicted protein of 589 aa with an alternative start codon encoded by exon 7. The first 7 amino acids are unique for this isoform. Isoforms 2-5b (620 aa) and 2-5c (631 aa) are nearly identical; both preserve the original translation initiation codon in exon 1, but differ in the splicing site between exons 13 and 20. This affects an alternative splice site (gc) at the end of exon 13, leading to the addition of 33 nucleotides to exon 13 (exon 13L) in case of the 2-5c transcript.

Since three novel classes of alternative *cRPGRIP1* transcripts were identified that involved inclusion of a 3'-alterna-

tive terminal exon 19a, 19c, or 19d, we focused on delineating the full structure of *cRPGRIP1* isoforms missing the RID domain. To date, we have characterized two of the new *cRPGRIP1* isoforms, 19c-1 and 19c-2 (GenBank accession numbers HQ315863 and HQ315864, respectively), that were found to be expressed in retinal tissue. Both isoforms are 3'-flanked by terminal exon 19c (Fig. 2). The 3' end of the 19c-1 and 19c-2 was obtained by 3'-RACE (with a gene-specific primer in exon 12) and 5' end was obtained by RLM 5'-RACE. The result indicates the existence of an alternative promoter (P2) located in intron 9, and TSS3 positionally coincides with the beginning of exon 10. The 19c-1 isoform arises from skipping exons 14 to 19 (encodes predicted protein of 212 aa), and 19c-2 lacks exons 14 to 18 (187 aa); the 3'UTR of 19c-2 includes part of exons 19 and 19c spliced together.

Based on our findings, we propose that the choice of 3' terminal exon can be specified by an upstream promoter. This putative mechanism may efficiently allow the synergistic function of 5' and 3'UTRs.

Tissue Distribution of *cRPGRIP1* Exon Combinations

The expression pattern of *cRPGRIP1* isoforms was examined by RT-PCR in mRNA isolated from retina, testis, heart, lung, and three brain regions (occipital, frontal, and temporal lobes). Nine primer pairs were design for different exon–exon combinations (Table 1). The results showed that most *cRPGRIP1* splice variants are predominantly expressed in retina at all ages tested (Fig. 3). Analysis of lung cDNA revealed that all splice variants tested appeared to be expressed, although the expression level of P1 promoter-driven isoforms 2-5a, 2-5b, and 2-5c was greatly reduced relative to that observed in the retina. Both 19c-1 and 19c-2 were also expressed in lung, where the expression level of 19c-1 was comparable to that in retina. We found a specific expression pattern of the *cRPGRIP1* isoforms tested in different tissues. For example, isoforms containing exons 1A, 1, and 2 are observed in all samples except for the temporal lobe. Analysis of exons 1A-2EX yields two splice variants in the retina, the lower molecular weight product arising from the skipping of exon 1. In addition to retina, splice variant skipping exon 1 was observed in heart and lung, but was not detected in testis or brain. The same expression pattern was observed for isoforms containing exon 3 and

FIGURE 2. Structure of the *cRPGRIP1* isoforms 2-5a, 2-5b, 2-5c, 19c-1, and 19c-2. The transcription (arrow shown at +1 on all transcripts) and predicted translation (arrow with Met indicated for each transcript) initiation sites are noted. The SMC/CC and RID domains of *RPGRIP1* encoded by the isoforms 2-5a, 2-5b, and 2-5c are represented. Note the partial RID for the 2-5b and 2-5c isoforms. No SMC/CC domain was present on 2-5a.

FIGURE 3. Analysis of tissue distribution of exon combinations of *CRPGRIP1* by RT-PCR. All isoforms were expressed in retina at all ages examined, and all tissues expressed isoforms containing 19 to 21. w, weeks; m, months; OL, occipital lobe; FL, frontal lobe; TL, temporal lobe.

exons 14 to 16. Expression of exons 19 to 21 of *CRPGRIP1* was observed in all tissues analyzed. An exonic splice combination specific for 19c-1 was found in testis, lung, and temporal lobe, but at lower levels than in retina. 19c-2, on the other hand, was present in testis and heart at levels comparable with that in retina, but expression in lung and temporal lobe was considerably lower.

Absolute Quantification of *RPGRIP1* Transcripts in Retina

To further evaluate the abundance of the identified *RPGRIP1* transcripts (reference transcript, isoforms 2-5a, 2-5b, 2-5c, 19c-1 and 19c-2), we used quantitative real-time RT-PCR of 6-month-old retina. Primer sets used in the analysis are shown in Table 1. First, the reference *RPGRIP1* transcript was evaluated and compared by using the two amplicons 1-3EX and 14-16EX. The standard curves for each fragment were comparable (Supplementary Fig. S1, <http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10-6094/-DCSupplemental>); thus, 14-16EX was subsequently used to characterize the reference sequence. Absolute quantification of the six selected tran-

scripts was performed by SYBR green real-time RT-PCR (Fig. 4). The lowest level of expression was found for the 2-5c (40 ± 7 molecules per sample) and 19c-1 (75 ± 3) transcripts, whereas the expression of the 2-5b and 19c-2 was found to be in several times higher (290 ± 23 and 265 ± 17 molecules per sample, respectively). The mean concentration of 2-5a was 3183 ± 114 molecules per sample, and the reference transcript concentration determined with the 14-16EX amplicon was highest at 21500 ± 1092 molecules per sample.

Contextual Analysis of *CRPGRIP1* Promoter Regions

To locate the position of the bovine, human, and murine putative promoters corresponding to canine P1, a comparative genomics approach was used to analyze a region 1000 bp upstream to 100 bp downstream of the *CRPGRIP1*'s first non-coding exon. The best promoter sequence similarity was found between -1 and -500 from the putative TSS1 position and consisted of 82.5% sequence identity between canine/bovine, 64.9% between canine/human, and 50.2% between canine/murine.

Using matrix match analysis (Launcher software GEMS; Genomatix GmbH, Munich, Germany), we used these alignments to identify potential regulatory motifs within the promoter region of the canine, bovine, human, and murine *RPGRIP1* orthologs. The approach identified potential promoter elements based on the defined orientation and distance correlated organization (framework) of a series of transcription factor (TF) binding sites. This in silico analysis revealed one common promoter framework in canine, bovine, and human promoter sequences consisting of a combination of four TF binding sites (Fig. 5). Three sites are positionally and structurally conserved motifs for CRX, PAX6, and FoxF2, suggesting that these TFs have potential roles in *RPGRIP1* regulation. PHOX2A, PHOX2, or CART1 motifs in the canine, human, and bovine sequences, respectively, were found adjacent to the PAX6 and FoxF2 motifs, suggesting a potential composite regulatory element in which PAX6 and FoxF2 motifs are separated by a 5-bp spacer from each other, and the FoxF2 motif overlaps with the PHOX2A/2 or CART1 motifs. Depending on the species, CRX is located 115 to 118 bp upstream of this cluster (Table 3). In comparison to the other three species, only partial similarity to the regulatory cluster described above was found in the murine putative promoter region (Fig. 5). The PAX6 and PHOX2 motifs were located positionally close to those described in other species, but putative minus-strand CRX binding site was found in position $-37/-57$. In addition, while the NR2E3 motif was identified in the canine ($-455/-431$), bovine ($-479/-455$), and human ($-258/-234$) sequences, it was not present in mouse.

Canine P2, flanking the *CRPGRIP1* coding exon 10 region, was analyzed in the same way. The highest promoter sequence similarity was found between -1 and -365 from the putative TSS3 position and consisted of 82.5% sequence identity between canine/bovine, 76.7% between canine/human, and 66.9% between canine/murine. No common complex regulatory modules were recognized in the orthologous promoter regions (FrameWorker algorithm; Genomatix, GmbH). However, the CRX motif was identified in bovine ($-354/-338$) and human ($-340/-324$) putative promoter sequences by a matrix match search (MatInspector algorithm; Genomatix GmbH). In the canine sequence, a CRX-like binding site, recognized as the RX motif, was found in position $-357/-341$ from TSS3. A putative minus-strand CRX-binding site was found in position $-252/-268$ in the murine sequence (Table 3).

FIGURE 4. Expression of *cRPGRIP1* reference and five alternatively spliced isoforms by absolute quantitative real-time RT-PCR. The copy number was highest for the 14-16EX amplicon, representing the reference transcript and lowest for isoforms 2-5c and 19c-1. Calibration curve is represented by open squares, and the sample analyzed is marked by black squares. Ct, cycle threshold.

DISCUSSION

Alternative splicing is an intrinsic regulatory mechanism of all metazoans that enables a single gene to code for multiple

FIGURE 5. Schematic representation of transcription factors binding sites determined by an algorithm that identified a common framework of elements (Frameworker; Genomatix) in the P1 promoter sequences of *RPGRIP1* orthologs (cf., *Canis familiaris*; Bt, *Bos taurus*; Hs, *Homo sapiens*; Mm, *Mus musculus*). The TSSs (arrow) are shown at +1 on all sequences. Note that there is one common promoter framework consisting of a combination of four TF binding sites in the canine, bovine, and human sequences. Three sites are positionally and structurally conserved motifs for CRX, PAX6, and FoxF2. PHOX2A, PHOX2, or CART1 motifs in the canine, human, and bovine sequences, respectively, were found adjacent to the PAX6 and FoxF2 motifs. In comparison to the other three species, the putative murine promoter region shows only partial similarity to the regulatory cluster.

proteins. In this article, we provide new information on the structure, organization, and expression of the *cRPGRIP1* reference sequence and five alternatively spliced variants: 2-5a, 2-5b, 2-5c, 19c-1, and 19c-2.

The *cRPGRIP1* reference transcript consists of 25 (24 coding) exons, and predicts a 136.07-kDa protein. The sequence exhibits high sequence homology with human (80.1% nt, 72.4% aa), bovine (81.5% nt, 73.1% aa), and murine (71.8% nt, 63.2% aa) orthologous reference sequences. Such high homology suggests a strong conservation of protein-domain organization, as has been shown for *RPGRIP1*.⁸⁻¹⁰ The N-terminal sequence of *cRPGRIP1* indicates the existence of a nuclear localization signal (NLS) motif for nuclear import of the nuclear domain (ND) of *RPGRIP1*.¹⁹ This sequence, K50RAR, is also conserved in the human (K72RLR) and bovine (K50RMR) proteins. A putative nuclear-localization sequence (NLS) R(71)RRR was found in the murine protein. A bipartite NLS was described in the RID of *RPGRIP1*,¹⁰ but analysis of the 3'-part of the *cRPGRIP1* sequence in the NLS database did not show potential motifs for nuclear transport in this region. Direct analysis of the canine *RPGRIP1* sequence with the protein BLAST tool identified two structurally different regions that were described for orthologous sequences: the SMC/CC domain encoded by exons 6 to 13 and the C2-domain encoded by exon 16 (Fig. 1A) (BLAST is provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD; www.ncbi.nlm.nih.gov/blast/).

In the present study, we describe the full structure of three alternatively spliced *cRPGRIP1* 5' products that share, together with the reference sequence, the same 3'-terminal exon 24, and are driven by a common promoter. In comparison with the reference, the 2-5a isoform encodes a predicted protein (65.7-kDa) with a truncated amino terminus (the first codon is

TABLE 3. Summary of Transcription Factor Binding Sites Determined by Database Search Algorithms in Promoter Sequences of *RPGRIP1* Orthologous Genes*

Transcription Factor	Family/Matrix	Species	Strand	Position from/ to TSS	Matrix Sim.	Sequence†
P1						
CRX (cone-rod homeobox)	V\$BCDF/V\$CRX.01	Cf	+	−215/−199	0.945	tctat TAAT c tcaaatc
		Bt	+	−218/−202	0.940	tctgt TAAT c tcagatc
		Hs	+	−218/−202	0.954	tcttt TAAT c tcaaatc
		Mm	−	−37/−53	0.981	gacac TAAT c ccaggcct
PAX6 (paired box 6)	V\$PAX6/V\$PAX6.03	Cf	−	−66/−84	0.773	tatt AACT C ttgtagatc
	V\$PAX6.03	Bt	−	−66/−84	0.767	tggt AACT C ttgtagatc
		Hs	−	−64/−82	0.770	tggt AACT C ttgtagatc
		Mm	−	−60/−78	0.767	tggt AACT C ttgtagctc
FoxF2 (forkhead box F2)	V\$FKHD/V\$FREAC2.01	Cf	+	−61/−45	0.898	aagagc TAAACA gaact
		Bt	+	−61/−45	0.898	aagagc TAAACA gaatt
		Hs	+	−59/−43	0.898	aaaagc TAAACA gaact
		Mm	−	−	−	−
PHOX2A (paired-like homeobox 2a)	V\$CART/V\$PHOX2A.01	Cf	−	−31/−50	0.841	aagagg AATTa atacagttct
CART1 (Cartilage homeoprotein 1)	V\$CART/V\$CART1.01	Bt	−	−34/−54	0.864	ggagt TAAT acaattctgttt
PHOX2	V\$CART/V\$PHOX2.01	Hs	−	−32/−52	0.893	ggagc TAAT ccagttctgttt
		Mm	−	−37/−57	0.934	gacac TAAT ccaggcctttta
NR2E3 (nuclear receptor subfamily 2, group E, member 3)	V\$NR2F/V\$PNR.01	Cf	+	−455/−431	0.843	atctaattt TCAAa atcaagttct
		Bt	+	−479/−455	0.843	atctgattt TCAAa atcaagttct
		Hs	+	−258/−234	0.801	ataaaagat TAAa gatcatcctct
		Mm	−	−	−	−
P2						
RX (retinal homeobox protein Rx)	V\$BCDF/V\$PCE1.01	Cf	+	−357/−341	0.900	atttt TAAT t cagcagg
CRX (cone-rod homeobox)	V\$BCDF/V\$CRX.01	Bt	+	−354/−338	0.974	atttc TAAT c caacagg
		Hs	+	−340/−324	0.972	atttc TAAT c cggcagg
		Mm	−	−252/−268	0.979	cacct TAAT c atgcatt

Cf, *Canis familiaris*; Bt, *Bos taurus*; Hs, *Homo sapiens*; Mm, *Mus musculus*; promoter 1, P1; promoter 2, P2; Sim., similarity.

* FrameWorker and MatInspector algorithms (Genomatix, Munich, Germany).

† Bold italic: consensus index value >60; capital letters: core sequence.

positioned in exon 7 and exons 8, 9, and 14 to 16 are skipped), which would indicate that the SMC/CC domain (encoded by exons 6 to 13^{8–10}) is not functional, and the C2 domain is absent. The absence of SMC/CC is a common feature of a range of short *RPGRIP1* isoforms, and some isoforms have been found in other species without C2 domains.^{9,18,22} The intact RID suggests that this isoform is involved in RPGR interactions. On the other hand, the 2-5b and 2-5c isoforms encode predicted proteins (67.7- and 70.8-kDa) that have the N-terminal SMC/CC domain, truncated RID, and no C2 domains. Both 2-5b and 2-5c have not previously been described in orthologous genes; thus, their physiologic function remains undetermined.

In addition, we described the full structure of two alternatively spliced *cRPGRIP1* products that have very short 5'UTR (1 nucleotide) and share the same 3'-terminal exon 19c and alternative TSS3, which coincides with the beginning of coding exon 10. The 19c-1 and 19c-2 isoforms encode predicted proteins of 23.76- and 20.87-kDa that lack RID. Analysis of protein sequences has not revealed any conservative protein domains and no analogues for those isoforms have been described in orthologous genes. It is interesting that in some short *RPGRIP1* mRNAs expressed from orthologous genes, for example human AK301780 and BC039089, the first codon is also positioned in exon 10. They have variably sized 5'UTR, 80 and 65 nucleotides (although 5'UTR can be incomplete), respectively, and are flanked by 3'-terminal exon 24. Together, these data suggest a complex control of transcription initiation of *RPGRIP1* isoforms as a subject worthy of further investigation. If, as shown herein, some *cRPGRIP1* isoforms are driven by an internal alternative promoter, this may be important for understanding the disparity in genotype–phenotype correlation observed in canine cone-rod dystrophy caused by a 44-nucleo-

tide insertion in exon 2 resulting in a premature stop.^{6,7} Those isoforms, with internal TSS, would potentially be expressed independent of the Ins44 genotype.

cRPGRIP1 pre-mRNA produces multiple alternatively spliced mRNA transcripts, but it is not known whether all these isoforms express protein products in the retina. For example, published Western blot data with antibodies against RID of *RPGRIP1* detected only two highly abundant translated products in the dog retina (~45- and 33-kDa), and a higher molecular mass protein ~160-kDa was observed only on longer exposure.⁸ This higher molecular mass protein most likely corresponds to the *cRPGRIP1* reference full-length product with a predicted size of 136.07-kDa, since not only size but charge as well contribute to migration of proteins in SDS-PAGE.

In this study, we provide new data for the structure and organization of the 5'- and 3'-ends of *cRPGRIP1*. To date, sequence analysis of the 5'-UTR of *RPGRIP1* transcripts has not been fully documented. It is known that the 5'-UTR of eukaryotic mRNAs plays a crucial role in the posttranscriptional regulation of gene expression through the modulation of translation efficiency, message stability, and subcellular localization.^{23–25} Our results indicate a complex organization of the 5'-UTR for the *RPGRIP1* gene. Using Cap-selective 5'-RACE, bioinformatics and RT-PCR analyses, we identified a novel 5'-splicing pattern for the *cRPGRIP1* which demonstrated a more complex structure than previously reported. In contrast to the 5' regulatory region, the 3'-UTRs of *cRPGRIP1* show low sequence homology with sequences of orthologous genes. We found that the 3'-UTR contains four polyadenylation sites, which define the border for short 3'-UTRs flanking the 3' terminal exons 19a (57 bp), 19c (34 bp), 19d (157 bp), and 24 (78 bp). Last, we cannot exclude the existence of additional

alternative PASs that contribute to thus far uncharacterized transcripts. Typically, 3'-UTR lengths vary between 21 and 8555 bp in humans and between 37 and 3324 bp in other mammals.²⁶ In general, polyadenylation events favor short 3'-UTRs, possibly to allow their expression without binding to regulatory factors quickly expressed within the cell. In comparison, longer UTRs may contain the regulatory motifs necessary to specify complex temporal and spatial translational programs in complex cells.^{26,27}

To evaluate the expression of *cRPGRIP1* splice variants containing specific exon-exon combination, we used RT-PCR. Expression of *cRPGRIP1* appears to be regulated in a highly tissue-specific manner. Exons 19, 20, and 21, part of sequence encoding RID, were found expressed in all tissues, but exon combination 13 to 19c was not detected in heart or occipital and frontal lobes of the brain. Although expression of exon-exon combinations specific for the 2-5a, 2-5b, and 2-5c isoforms were also observed at low levels in the lung, their relatively high expression in the retina of all ages tested indicates an important role for these in photoreceptor function. Of note, exon 1A of *cRPGRIP1* is also expressed in testis, heart, lung, and brain, suggesting that the tissue-specific promoter is either located positionally upstream of the retinal *RPGRIP1* promoter P1 or is governed by local transcription factors.

Notably, we observed a significant difference in expression level between the reference and the five spliced variants of *cRPGRIP1* within the retina, an indication that regulation of isoform expression is at the transcriptional level, or depends on isoform stability. These results suggest that transcription of alternatively spliced isoforms of *RPGRIP1* may be regulated by exonic and/or intronic regulatory elements (for review, see Ref. 28). It is interesting that the only difference between isoforms 2-5b and 2-5c is a longer exon 13 (exon 13L with 33 additional nucleotides), yet the difference results in an approximately sevenfold lower expression level for 2-5c.

RPGRIP1-knockout mice^{16,17} and dogs with naturally occurring disease⁶ illustrate the importance of *RPGRIP1* for retina function and development. Although such data suggest that early expression of *RPGRIP1* is essential during retinal development, the underlying mechanism of transcriptional regulation of *RPGRIP1* remains unknown. By contextual analysis of the *RPGRIP1* promoter region P1, we demonstrated that this region has a structurally conserved regulatory module between orthologous canine, human, and bovine gene sequences. Arrangement of the four motifs, three of which—PAX6, FoxF2, and Phox2A (or CART1)—form a potential composite element that could be regulated by the assemblage of several alternative complexes of transcription factors on the *RPGRIP1* promoter under various conditions. Although the role of FoxF2 and Phox2A in retinal development is unclear, they are important developmental regulators in other tissues.^{29–31} PAX6 expression by progenitor cells is essential to begin developmental processes in the eye and retina,^{32,33} and is expressed in all proliferating retinal neuroblasts until they exit the cell cycle. At that point, photoreceptors stop expressing PAX6 and begin expressing CRX.^{32,34} Sequence analysis of the putative promoter region P2 has shown its more simple organization in comparison to P1. CRX was suggested as the main transcription factor controlling expression of the *RPGRIP1* isoforms that are driven by an internal promoter P2. How these transcription factors regulate *RPGRIP1* expression in development and disease is still unknown and will be the subject of future investigations.

In conclusion, the results reported herein provide new insights into the structure and organization of the canine *RPGRIP1* gene. It sets the stage for future studies to examine how *cRPGRIP1* transcript heterogeneity relates to its expres-

sion at the protein level and to evaluate how altered gene expression contributes to disease.

Acknowledgments

The authors thank Sem Genini for providing canine cDNA samples from retinas of 3- and 16-week-old dogs and Mary Leonard for illustrations.

References

1. Dryja TP, Adams SM, Grimsby JL, et al. Null *RPGRIP1* alleles in patients with leber congenital amaurosis. *Am J Hum Genet.* 2001; 68:1295–1298.
2. Wiik AC, Wade C, Biagi T, et al. A deletion in nephronophthisis 4 (NPHP4) is associated with recessive cone-rod dystrophy in standard wire-haired dachshund. *Genome Res.* 2008;18:1415–1421.
3. Gerber S, Perrault I, Hanein S, et al. Complete exon-intron structure of the RPGR-interacting protein (RPGRIP1) gene allows the identification of mutations underlying leber congenital amaurosis. *Eur J Hum Genet.* 2001;9:561–571.
4. Booi JC, Florijn RJ, ten Brink JB, et al. Identification of mutations in the *AIPL1*, *CRB1*, *GUCY2D*, *RPE65*, and *RPGRIP1* genes in patients with juvenile retinitis pigmentosa. *J Med Genet.* 2005;42: e67.
5. Hameed A, Abid A, Aziz A, Ismail M, Mehdi SQ, Khaliq S. Evidence of *RPGRIP1* gene mutations associated with recessive cone-rod dystrophy. *J Med Genet.* 2003;40:616–619.
6. Mellersh CS, Boursnell ME, Pettitt L, et al. Canine *RPGRIP1* mutation establishes cone-rod dystrophy in miniature longhaired dachshunds as a homologue of human leber congenital amaurosis. *Genomics.* 2006;88:293–301.
7. Miyadera K, Kato K, Aguirre-Hernandez J, et al. Phenotypic variation and genotype-phenotype discordance in canine cone-rod dystrophy with an *RPGRIP1* mutation. *Mol Vis.* 2009;15:2287–2305.
8. Castagnet P, Mavlyutov T, Cai Y, Zhong F, Ferreira P. *RPGRIP1*s with distinct neuronal localization and biochemical properties associate selectively with RanBP2 in amacrine neurons. *Hum Mol Genet.* 2003;12:1847–1863.
9. Roepman R, Bernoud-Hubac N, Schick DE, et al. The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Hum Mol Genet.* 2000;9:2095–2105.
10. Roepman R, Letteboer SJ, Arts HH, et al. Interaction of nephrocystin-4 and *RPGRIP1* is disrupted by nephronophthisis or leber congenital amaurosis-associated mutations. *Proc Natl Acad Sci USA.* 2005;102:18520–18525.
11. Mavlyutov TA, Zhao H, Ferreira PA. Species-specific subcellular localization of RPGR and *RPGRIP1* isoforms: Implications for the phenotypic variability of congenital retinopathies among species. *Hum Mol Genet.* 2002;11:1899–1907.
12. Boylan JP, Wright AF. Identification of a novel protein interacting with RPGR. *Hum Mol Genet.* 2000;9:2085–2093.
13. Hong DH, Yue G, Adamian M, Li T. Retinitis pigmentosa GTPase regulator (RPGR)-interacting protein is stably associated with the photoreceptor ciliary axoneme and anchors RPGR to the connecting cilium. *J Bio Chem.* 2001;276:12091–12099.
14. Lu X, Ferreira PA. Identification of novel murine- and human-specific *RPGRIP1* splice variants with distinct expression profiles and subcellular localization. *Invest Ophthalmol Vis Sci.* 2005;46: 1882–1890.
15. Shu X, Fry AM, Tulloch B, et al. RPGR ORF15 isoform co-localizes with *RPGRIP1* at centrioles and basal bodies and interacts with nucleophosmin. *Hum Mol Genet.* 2005;14:1183–1197.
16. Zhao Y, Hong DH, Pawlyk B, et al. The retinitis pigmentosa GTPase regulator (RPGR)-interacting protein: subserving RPGR function and participating in disk morphogenesis. *Proc Natl Acad Sci USA.* 2003;100:3965–3970.
17. Won J, Gifford E, Smith RS, et al. *RPGRIP1* is essential for normal rod photoreceptor outer segment elaboration and morphogenesis. *Hum Mol Genet.* 2009;18:4329–4339.

18. Ferreira PA. Insights into X-linked retinitis pigmentosa type 3, allied diseases and underlying pathomechanisms. *Hum Mol Genet.* 2005;14 Spec No. 2:R259-R267.
19. Lu X, Guruju M, Oswald J, Ferreira PA. Limited proteolysis differentially modulates the stability and subcellular localization of domains of RPGRIP1 that are distinctly affected by mutations in leber's congenital amaurosis. *Hum Mol Genet.* 2005;14:1327-1340.
20. Lheriteau E, Libeau L, Stieger K, et al. The RPGRIP1-deficient dog, a promising canine model for gene therapy. *Mol Vis.* 2009;15:349-361.
21. Zangerl B, Sun Q, Pillardy J, et al. Development and characterization of a normalized canine retinal cDNA library for genomic and expression studies. *Invest Ophthalmol Vis Sci.* 2006;47:2632-2638.
22. Strausberg RL, Feingold EA, Grouse LH, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci USA.* 2002;99:16899-16903.
23. Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. *Genome Biol.* 2002;3:REVIEWS0004.
24. Pickering BM, Willis AE. The implications of structured 5' untranslated regions on translation and disease. *Semin Cell Dev Biol.* 2005;16:39-47.
25. Payton SG, Haska CL, Flatley RM, Ge Y, Matherly LH. Effects of 5' untranslated region diversity on the posttranscriptional regulation of the human reduced folate carrier. *Biochim Biophys Acta.* 2007;1769:131-138.
26. Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene.* 2001;276:73-81.
27. Doran G. The short and the long of UTRs. *J RNAi Gene Silencing.* 2008;4:264-265.
28. Wang Z, Burge CB. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA.* 2008;14:802-813.
29. Paris M, Wang WH, Shin MH, Franklin DS, Andrisani OM. Homeodomain transcription factor Phox2a, via cyclic AMP-mediated activation, induces p27Kip1 transcription, coordinating neural progenitor cell cycle exit and differentiation. *Mol Cell Biol.* 2006;26:8826-8839.
30. Ormestad M, Astorga J, Landgren H, et al. Foxf1 and Foxf2 control murine gut development by limiting mesenchymal wnt signaling and promoting extracellular matrix production. *Development.* 2006;133:833-843.
31. Aitola M, Carlsson P, Mahlapuu M, Enerback S, Pelto-Huikko M. Forkhead transcription factor FoxF2 is expressed in mesodermal tissues involved in epithelio-mesenchymal interactions. *Dev Dyn.* 2000;218:136-149.
32. Garelli A, Rotstein NP, Politi LE. Docosahexaenoic acid promotes photoreceptor differentiation without altering crx expression. *Invest Ophthalmol Vis Sci.* 2006;47:3017-3027.
33. Hsieh YW, Yang XJ. Dynamic Pax6 expression during the neurogenic cell cycle influences proliferation and cell fate choices of retinal progenitors. *Neural Dev.* 2009;4:32.
34. Hennig AK, Peng GH, Chen S. Regulation of photoreceptor gene expression by crx-associated transcription factor network. *Brain Res.* 2008;1192:114-133.