Genomics 96 (2010) 362–368

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Exonic SINE insertion in *STK38L* causes canine early retinal degeneration (*erd*)

Orly Goldstein [a,1], Anna V. Kukekova [a,1], Gustavo D. Aguirre [b], Gregory M. Acland [a,*]

[a] Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Hungerford Hill Road, Ithaca NY 14853, USA
[b] Department of Clinical Studies, School of Veterinary Medicine, University of Pennsylvania, 3900 Delancey St., Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 16 July 2010
Accepted 22 September 2010
Available online 29 September 2010

Keywords:
Retinal degeneration
Leber Congenital Amaurosis
STK38L
Animal model

ABSTRACT

Fine mapping followed by candidate gene analysis of *erd* — a canine hereditary retinal degeneration characterized by aberrant photoreceptor development — established that the disease cosegregates with a SINE insertion in exon 4 of the canine *STK38L/NDR2* gene. The mutation removes exon 4 from *STK38L* transcripts and is predicted to remove much of the N terminus from the translated protein, including binding sites for S100B and Mob proteins, part of the protein kinase domain, and a Thr-75 residue critical for autophosphorylation. Although known to have roles in neuronal cell function, the STK38L pathway has not previously been implicated in normal or abnormal photoreceptor development. Loss of *STK38L* function in *erd* provides novel potential insights into the role of the STK38L pathway in neuronal and photoreceptor cell function, and suggests that genes in this pathway need to be considered as candidate genes for hereditary retinal degenerations.

1. Introduction

Early retinal degeneration (*erd*) is a canine early onset autosomal recessive disease [1] corresponding to Leber Congenital Amaurosis in humans. It is characterized by aberrant functional and structural development of rod photoreceptor inner and outer segments, and of both rod and cone synapses, with subsequent retinal degeneration [1]. The locus was previously mapped to a 30 cM interval on canine chromosome 27 (CFA27), corresponding to a part of human chromosome 12p [2].

In this report, the candidate region was first reduced by fine-scale mapping to a 4.6 cM zero recombination interval on CFA27, a 1.6 Mb region in the canine genome sequence assembly (CanFam2). Positional candidate gene analysis then identified a tRNA-like short interspersed repeat (SINE) insertional mutation in exon 4 of *STK38L* that cosegregates with *erd*.

The gene *STK38L*, also called *NDR2*, codes for serine/threonine kinase 38-like protein — a protein kinase in the nuclear Dbf2-related (NDR) family involved in neuronal cell division and morphology [3–5]. STK38L kinase is activated by S100B (a Ca^{++}-binding protein) and Mob (Mps one binder), enabling STK38L to autophosphorylate at Thr-75, Ser-282, and Thr-444 [5,6]. STK38L and its partners have been implicated in neuronal cytoskeletal development, neurite outgrowth and synaptic remodeling. Despite this, much remains to be determined concerning their exact role and mechanisms of action. A particular gap in this understanding is their role in the retina. Until now, neither *STK38L* nor its partners, has been implicated in any hereditary retinal degeneration, or as having critical roles in photoreceptor development. The *erd*-mutant dog therefore offers potential insights into the role of this pathway in both normal and abnormal photoreceptor and neuronal development.

2. Material and methods

2.1. Animals

The *erd* strain of dogs, maintained at the Retinal Disease Studies Facility (RDSF) in Kennett Square, PA, derives from 2 *erd*-affected Norwegian elkhound dogs, highly inbred siblings, that were outbred to Beagles and other dogs, and their progeny used to generate pedigrees segregating *erd*. Selected dogs from this strain were evaluated by DNA and RNA analysis. Blood samples were also collected from privately owned dogs of several breeds, and DNA extracted as previously described [7]. All procedures involving animals were undertaken according to IACUC approved protocols at Cornell University and the University of Pennsylvania, and in adherence to the ARVO Resolution for the Use of Animals in Ophthalmic and Vision Research.

2.2. Gene-associated markers for linkage mapping

When fine mapping of *erd* began, the canine genome assembly was not available. The *erd* locus had previously been mapped to the CFA27 interval flanked by genes PFKM, LALBA, PTHLH, and IAPP [2], homologously located on HSA12p. From comparison of the canine and human genomes [8], 7 genes (*BICD1*, *PTHLH*, *ITPR2*, *SSPN*, *SHARP1*,

* Corresponding author. Fax: +1 607 256 5608.
E-mail address: gma2@cornell.edu (G.M. Acland).
[1] These 2 authors contributed equally to the work described in this manuscript.

0888-7543/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2010.09.003

LRMP, and *SIAT8A*) were selected from HSA12p between *LALBA* and *IAPP*, for fine mapping *erd*.

The RPC181 canine 8.1-fold BAC library from Roswell Park Cancer Institute (http:www.chori.org/bacpac/mcanine81.htm) was screened to identify BAC clones specific for *SIAT8A*, *BICD1*, and *SHARP1* as previously described [9]. Gene-associated polymorphic markers were developed from identified BAC clones (Supplementary Table 1A) as previously described [9].

For *ITPR2*, *PTHLH* and *SSPN*, partial canine genomic sequences were obtained from the Institute for Genome Research (TIGR, E. Kirkness, pers. comm.) and used to develop gene-specific polymorphic markers (Supplementary Table 1A). A polymorphic (CA)n repeat was identified in the *LRMP* gene by amplifying and sequencing canine intron 1 using primers designed from human sequences.

When the canine genome assembly (CanFam2) became available the position of all markers was identified in CanFam2 using UCSC Bioinformatics tools (electronic PCR, Blat; <http://genome.ucsc.edu/cgi-bin/hgGateway >).

2.3. Genotyping and linkage analyses

Pedigrees segregating *erd* were genotyped for informative markers (Supplementary Table 1). Markers were amplified from 50 ng of genomic DNA as per: 96 °C for 2 min; 30 cycles of 96 °C (20 s), 58 °C (20 s), and 72 °C (20 s); and a final extension at 72 °C for 5 min. PCR products were analyzed either in 12% polyacrylamide gel or on an ABI310 Genetic Analyzer (Applied Biosystems, Foster City, CA). For the latter method, forward primers had a 5′-end fluorescent label 6Fam, Hex, or Tet (GIBCO® Invitrogen, Carlsbad, CA).

Marker genotypes were checked for Mendelian segregation and analyzed to establish linkage and map order using MultiMap [10] as previously described [11]. A sex-averaged framework map was constructed with LOD 3.0 support starting with the most informative markers (*BICD1* and *SIAT8A*). Further markers were added with LOD 2.0 support (Fig. 1). Haplotype analysis of *erd* pedigrees was performed manually following "the minimal recombination number" rule.

2.4. CFA27 radiation hybrid mapping

To refine marker order on CFA27 and saturate the map with gene-specific markers, an RH Map of CFA27 was built using a commercially available 3000 rd canine radiation hybrid panel (RH3000) (Research Genetics, Huntsville, Al). The 22 markers selected for RH mapping, comprised 3 microsatellites linked meiotically to *erd* (FH2289, LEI002, REN277K09), 12 markers selected from the CFA27 RH Map 5000 cR [8], and 7 gene-specific markers (*BICD1, ITPR2, ARG99, SHARP1/BHLHE41, BCAT1*, *LRMP, SIAT8A/ST8SIA1*) derived from sequence of gene-specific canine BAC clones or the 3′UTR and intron sequences of the canine genes (Fig. 1; Supplementary Table 1A). Markers were amplified and an RH map constructed as described previously [9]. When the canine genome assembly CanFam2 became available, the position of all markers was identified using UCSC Bioinformatics tools (electronic PCR, Blat; <http://genome.ucsc.edu/cgi-bin/hgGateway >).

2.5. Candidate gene selection

Within the final LD region, 15 *Refseq* genes were identified. Initial candidate gene screening was limited to genes expressed in retina or eye (either mRNA or ESTs deposited in database) and mapped to HSA12p12.1-11.22, which exhibits conserved synteny with CFA27. The most proximal gene (*KLHDC5*) and the most distal two genes (*SSPN* and *ITPR2*) were also excluded from initial screening because of their proximity to the ends of the interval. As none of the remaining 7 genes had previously been implicated in an hereditary retinal degeneration, or otherwise identified as critical for photoreceptor development, all were considered equally viable candidates (Supplementary Tables 1B and Ci).

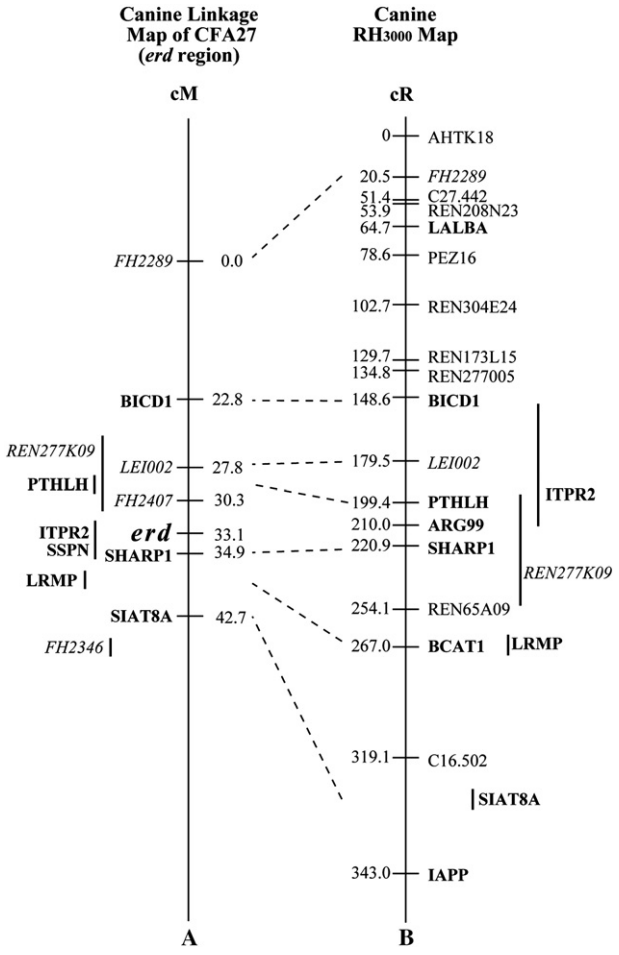

Fig. 1. Fine mapping of the *erd* interval on canine chromosome 27. A. Meiotic linkage map of the *erd* interval on canine chromosome 27 (CFA27) established by multipoint linkage analysis. The *erd* locus lies in a 4.6 cM zero recombination interval between microsatellite FH2407 and a gene-specific polymorphism in *SHARP1*. The gene-specific polymorphism in *PTHLH* cosegregates with FH2407. B. Radiation hybrid map of the corresponding interval. Marker order is in agreement with the linkage map (A) for markers mapped both ways. The *erd* interval corresponds to the region between *PTHLH* and *SHARP1*, a distance of 1.5 cR_{3000}.

2.6. PCR amplification and sequencing

Primers were designed to amplify coding sequences or genomic fragments within genes under standardized amplification conditions (Tm between 56 °C and 63 °C). Products were sequenced using the Applied Biosystems Automated 3730 DNA Analyzer (Applied Biosystems, Foster City, CA). Sequences were then analyzed and compared using Sequencher® 4.2.2 Software (Gene Codes Corporation, Ann Arbor, MI).

2.7. Population screening for the 4 bp deletion in intron 3

PCR screening of the 4-base deletion was done using three primers: two gene-specific primers flanking the 4 bp polymorphism (Supplementary Table 1Cii, primer pair 7) and a fluorescent-labeled FAM-M13(-21) universal primer: 5′-FAM-TGTAAAACGACGGCCAGT-3′. Unlabeled M13(-21) sequence (18 bp) was fused to the 5′ end of the forward primer. The reaction used: 2 pmol FAM-M13(-21) forward primer; 4 pmol reverse primer; and 1 pmol M13-tailed forward primer, made to a final volume of 11 µl with master mix (Invitrogen standard PCR reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 1 U of Taq

polymerase and 20 ng of DNA). PCR cycles included an initial denaturing step of 96 °C (2 min), 25 cycles of 95 °C (30 s), 56 °C (45 s), 72 °C (30 s), then 10 cycles of 95 °C (30 s), 53 °C (45 s), 72 °C (30 s) with a final extension step of 72 °C for 10 min. Subsequently, 1 μl of the PCR product was mixed with 0.2 μl of size marker Liz-500 (Applied Biosystems, Foster City, CA) and 9.8 μl of formamide and ran on the ABI 3730xl DNA Analyzer. Fragment size was analyzed using GeneMapper v3.0 (Applied Biosystems, Foster City, CA).

2.8. Population screening, exon 4 SINE insertion

To screen canine populations for this SINE insertion, primers flanking it were designed (Supplementary Table 1Cii, primer pair number 5). Reactions used GoTaq green master mix (Promega, Madison, WI), an initial denaturing step of 94 °C (10 min), 30 cycles of 94 °C (1 min), 58 °C (1 min), 72 °C (1.5 min), and a final extension step of 5 min at 72 °C. Products were visualized on 1.8% agarose gel: a normal dog showed one 625 bp band (homozygous for the normal allele), an affected dog showed one 854 bp band (homozygous for the insertional mutant allele), and a carrier dog had 2 bands representing both alleles, normal and mutant.

2.9. Tissue collection, RNA extraction, cDNA preparation, northern analysis

Tissue samples were collected as previously described [7]. Total RNA was extracted from frozen normal, carrier and affected tissues at 4.7, 8.7 and 14.1 weeks old using TRIZOL (Invitrogen, Carlsbad, CA) following the manufacturer's protocol, with 1 μl of RNase inhibitor (SUPERase-In, Ambion, Austin, TX) added to prevent RNA degradation. RNA was also extracted from a normal 8.7 week old dog from kidney, large intestine, small intestine, stomach, cerebellum, occipital cerebral cortex, frontal cerebral cortex, amygdala, hippocampus, heart, skeletal muscle and testis. Testis RNA from an 8.3 week affected dog was also extracted. Five to 10 μm of total RNA in 2 μl volume was used to prepare cDNA using Thermoscript RT-PCR kit (Invitrogen, Carlsbad, CA). Extension of cDNA sequence by 5′ and 3′ RACE was obtained through the SMART™RACE cDNA amplification kit (Clontech, Mountain View, CA) following the manufacturer's protocol (Supplementary Table 1Cii, primers 1 and 2). RNA membranes were prepared as previously described [7]. A probe for northern analysis was generated from normal retinal cDNA using gene-specific primers from the last exon of the gene, exon 14, which shows the weakest homology to *STK38* (*NDR1*) another member of the *STK* gene family (Supplementary Table 1Cii, primer pair 8). Probe was prepared, labeled, and hybridized to membrane as previously described [7]. Blots were exposed to X-ray film (Kodak, Rochester, NY) at −70 °C for 5 days with two intensifying screens (Kodak, Rochester, NY). The same membranes were subsequently hybridized to a canine beta-actin probe as a control for RNA loading, and exposed over night at room temperature.

2.10. In-silico analysis

The predicted canine STK38L sequence was aligned to human, murine, chicken and zebrafish STK38L proteins using MegAlign software (DNASTAR, Lasergene, Madison, WI). Protein structure and prediction of domains used the SWISS-PROT (http://us.expasy.org/sprot) and NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi) databases.

3. Results

3.1. Fine scale mapping of the erd interval.

Genotypes were obtained from 17 *erd*-segregating pedigrees, including 141 animals in informative generations, for 5 previously published canine microsatellite markers [8] and gene-specific microsatellite and SNP markers identified in seven genes: *BICD1*, *PTHLH*, *ITPR2*, *SSPN*, *SHARP1*, *LRMP*, *SIAT8A* (see Material and methods, Section 2.2, and Fig. 1). Multipoint analysis placed the *erd* locus and six markers at unique map positions at LOD 2 support, and the most likely location of other markers was assigned (Fig. 1). Two-point linkage analysis confirmed the tightest linkage between *erd* and markers SHARP1 (theta = 0.014; lod = 36.075) and FH2407 (theta = 0.028; lod = 25.355). From two-point linkage and haplotype analyses, markers PTHLH, FH2407, LEI002, and REN277K09 form a cluster with no internal recombinations observed, but sharing one recombination with *erd*. Similarly, one recombination was detected between *erd* and SHARP1. The position of the *erd* locus, its final absolute linkage disequilibrium (LD) interval, was thus refined to the 4.6 cM interval flanked by genes PTHLH and SHARP1, corresponding to the 1.6 Mb CanFam2 region of CFA27 between nucleotides 22,803,789 and 24,476,056.

The order of markers on CFA27 established by RH mapping was consistent with those by linkage mapping (Fig. 1) and in CanFam2 (Supplementary Table 1). Gene order in this RH3000 Map was also highly conserved compared to homologous regions of the human GB4 map and genome sequence assembly.

3.2. Evaluation of positional candidate genes

Of the 15 *Refseq* genes in the final LD interval for *erd* (Fig. 2A), the coding regions of 7 were amplified by RT-PCR and sequenced. *MRPS35* (accession # GU811152), *PPFIBP1* (GU811154; exons 1–7 only), *MED21* (GU811151), *TM7SF3* (HM013658), *FGFR1OP2* (GU811150), and *C12orf11* (GU811149), showed no size differences between products from affected and normal samples, and no significant sequence differences between cases and controls. Several polymorphisms were identified that were not associated with *erd* disease status (see accession deposits for details). Significant differences in size and sequence were detected for *STK38L* (GU811153).

3.3. Mutation identification in STK38L

3.3.1. Canine STK38L mRNA

BLAT alignment of the human *STK38L* mRNA (NM015000.3) — comprising 14 exons, 13 of which code for a 464 amino acid protein — against the CanFam2 genomic sequence, identified 14 predicted canine exons, with the first methionine located in the second exon (Fig. 2B), a structure shared with its human and murine orthologs. The predicted canine *STK38L/NDR2* gene and mRNA share about 78% identity with the predicted canine *STK38* paralogous gene (also known as *NDR1*, on CFA12). Two primer pairs (Supplementary Table 1Ci, primers 1–2) specific for *STK38L* successfully amplified the complete coding region in 2 overlapping fragments from retinal cDNA of a 10.4 week old normal dog (accession # GU811153, Supplementary Fig. 1A1). By 5′ RACE, (Supplementary Table 1Cii, primer 1) the cDNA sequence was extended 48 bases upstream to the first ATG located in exon 2. This result is similar to that in the homologous mouse sequence AY223819 [4]. Because the 5′ end of the sequence is open, further upstream extension was considered possible. However, *in-silico* analysis of the 5′ end of canine *STK38L* identified several transcription elements and signals that support the assumption that this is the end of the gene: three SP1 elements (gggcgg) are located upstream to the first ATG at positions −110 bp, −111 bp and −288 bp in exon 1. One CRE element is found in position −228, and a TATA box is present in position −987 bp.

3′ RACE using primer 2 in Supplementary Table 1Cii resulted in two products: one extended 265 nt beyond the stop codon in exon 14, and included a polyadenylation signal 28 bases preceding the poly A tail. The second extended 921 nt after the stop and included a polyadenylation signal 25 bases preceding the poly A tail (Supplementary Fig. 1A1).

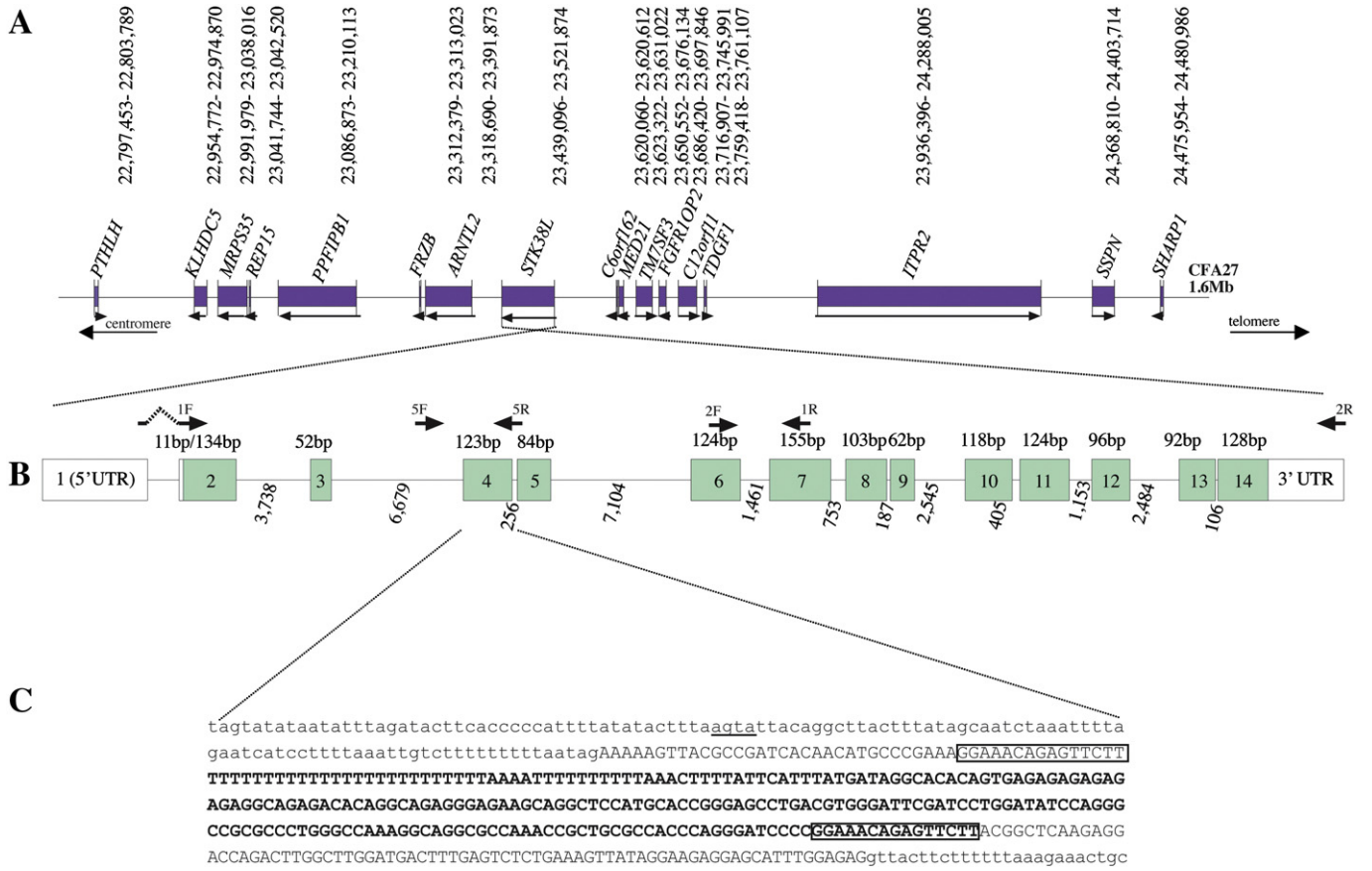

Fig. 2. Schematic illustration of the *erd* interval on canine chromosome 27. A. The 1.6 Mb zero recombination interval is drawn to scale and includes 15 Refseq genes between *PTHLH* and *SHARP1*. Their position on the UCSC CanFam2 dog assembly is indicated by nucleotide numbers, and their annotation by black arrows. B. Canine *STK38L* comprises 14 exons. Filled boxes represent coding regions and empty boxes the 5′ and 3′ untranslated regions. The sizes of the exons and introns are marked above and below the boxes, respectively. Primer pairs 1 and 2, represented by black arrows, were used to amplify the complete coding region from normal and affected dogs, while primer pair 5 was used to characterize the genomic region flanking exon 4. C. Structure of the inserted SINE. The nucleotide sequence of the tRNA-like SINE is depicted in bold upper case letters. Exon 4 sequence observed in non-affected dogs is indicated in unbold upper case letters. The 15 nucleotide sequence representing the direct duplication at the insertion site is boxed. Introns 3 and 4 are indicated in lower case letters. Underlined is the polymorphic 4 nucleotide sequence (agta) site deleted in *erd*-affected dogs and in a small number of normal dogs from other breeds.

3.3.2. Evaluation of STK38L in erd-affected dogs

Comparison of RT-PCR products from *erd*-affected, carrier and normal retina (9.6, 9.1 and 6.9 weeks old, respectively) showed no sizes differences when amplifying from exon 6 to 14 (Supplementary Table 1Ci, primer pair 2). However, when amplifying from exon 1 to 7 (Supplementary Table 1Ci, primer pair 1), the product size varied with *erd* genotype. The affected dog yielded an amplimer 123 nt shorter than that of the homozygous normal dog, and the heterozygote yielded both amplimers (Fig. 3A). Sequence analysis revealed that exon 4 was absent from the affected mRNA. No other differences were observed between affected and normal dogs (Supplementary Fig. 1B1).

To determine why exon 4 was absent, genomic donor–acceptor sites in introns 3 and 4 were sequenced from *erd*-affected, carrier and normal dogs (Supplementary Table 1Cii, primer pairs 3 to 6). Primer pair 5 amplified the expected 625 bp long product from the normal dog but an 854 bp product from affecteds, and both products from carriers. Sequence analysis showed a 4 nt deletion in intron 3 (−68 to −72) and insertion of a tRNA-like short interspersed repeat (SINE) in exon 4 including a 15 nt duplication (Fig. 2C). The 854 bp amplicon cosegregated with *erd* (Fig. 3B). Sequence analysis of 8 affected and 3 carrier colony dogs from generations closer to the proband showed concurrence of the SINE insertion and the 4 nt deletion, suggesting that neither mutation arose within the colony.

3.3.3. Population screening for the 4 nt deletion and the SINE insertion in STK38L

To establish the disease relevance of the 4 nt deletion in intron 3 and the SINE insertion in exon 4, we first screened a set of 82 purebred dogs from 16 different breeds that did not include Norwegian elkhounds for the 4 nt deletion. Amplified genomic DNA flanking the 4 nt deletion site established that 74 of these dogs were homozygous for the wild type allele (no deletion, 279 bp product), 6 were heterozygous for the deletion (yielding a 275 bp and a 279 bp product, 3 American Eskimos, 2 Border Collies and one American cocker spaniel) and 2 dogs were homozygous for the 4 nt deletion (one American cocker spaniel and one Border collie, Supplementary Table 2A). The latter two dogs were not affected with *erd* or any other form of retinal degeneration, and their amplicons were sequenced and confirmed to have the same 4 nt deletion. All 8 dogs with the 4 nt deletion were screened for the SINE insertion by amplification of genomic DNA flanking the insertion. No SINE insertion in exon 4 was found in these 8 dogs. These results suggest that the 4 nt deletion in intron 3 represents the minor allele of a polymorphism that does not cause *erd*, and that the *erd* mutation occurred on a chromosome bearing this minor allele.

A further set of control dogs (141 dogs from 24 breeds) were screened for the SINE insertion and all were homozygous wild type (Supplementary Table 2B).

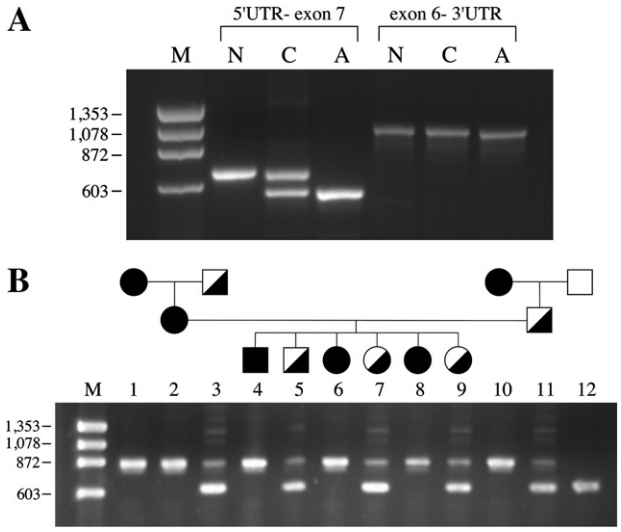

Fig. 3. A SINE insertion in *STK38L* cosegregates with disease in *erd*-affected dogs. A. Gel electrophoresis of the complete coding sequence of canine *STK38L*, amplified in two overlapping fragments (5′UTR-exon 7; exon 6–3′UTR) from normal (N), carrier (C), and affected (A) dogs. Primers located on the 5′UTR and exon 7 amplify a 661 bp amplicon from a normal dog, a 538 bp amplicon from an affected dog, and both bands from a heterozygous dog. No size differences among normal, affected and heterozygous products were observed for the amplification product including exons 6 to 14 (exon 6–3′UTR). B. Segregation of the SINE insertion in a 3-generation *erd*-informative pedigree. The larger band (854 bp) represents the affected amplicon which includes the SINE insertion; the 625 bp band represents the normal amplicon. All affected dogs (1,2,4,6,8,10) yield only the 854 bp band; the homozygous normal grandsire (12) yields only the 625 bp band; and the heterozygous dogs (3,5,7,9,11) yield both bands (and fainter heterodimer bands that run slowly). There is complete cosegregation between the SINE insertion and the disease.

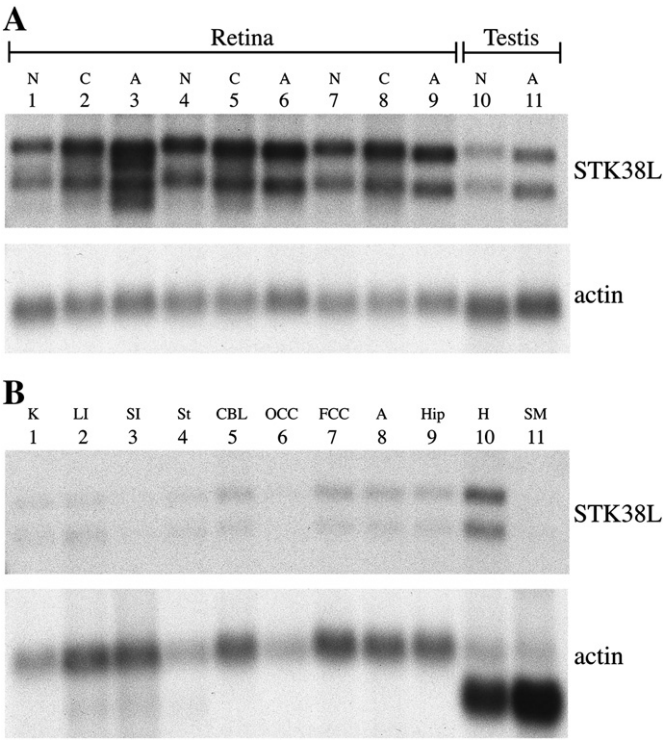

Fig. 4. Northern analysis of canine *STK38L* expression in retina and testis. A. *STK38L* expression in normal [N], carrier [C] and affected [A] canine retinas at 3 time points [lanes 1–3 ~4.5 weeks, lanes 4–6 ~8.5 weeks, lanes 7–9 ~14 weeks]; and in normal [N] and affected [A] testis at ~8.5 weeks old [lanes 10–11]. Expression of *ß-actin* is shown for evaluation of loading among lanes. Two STK38L transcripts (~4.4 kb and ~5.1 kb) are observed in normal retina and testis. The ~4.5 week old affected retina (lane 3) demonstrates a different pattern with 4 "shadow-like" bands, but this pattern was not observed in affected retina at later ages (lanes 6 and 9). B. *STK38L* expression profile in kidney (K, lane 1), large intestine (LI, lane 2), small intestine (SI, lane 3), stomach (St, lane 4), cerebellum (CBL, lane 5), occipital cerebral cortex (OCC, lane 6), frontal cerebral cortex (FCC, lane 7), amygdala (A, lane 8), hippocampus (Hip, lane 9), heart (H, lane 10), and skeletal muscle (SM, lane 11) from an 8.7 week old normal dog. Expression of *ß-actin* is shown for evaluation of loading among lanes. *STK38L* is expressed strongly in heart (lane 10), at much lower levels in kidney, large intestine, small intestine, stomach and brain (lanes 1–9) and is not detected in skeletal muscle (lane 11).

Seven Norwegian elkhound dogs segregating a retinal degeneration clinically distinguishable from *erd* were also tested for both the SINE insertion and the 4 nt deletion; all were homozygous wildtype by both tests, confirming that this breed has more than one form of inherited retinal degeneration segregating.

3.4. RNA expression analysis

Northern analysis using an *STK38L* specific 3′ end probe (see Material and methods, Section 2.9) identified 2 transcripts (4.4 kb and 5.1 kb) in retina from normal and carrier dogs at ages ranging from 4 to 14 weeks old (Fig. 4A). In the 4.3 week old affected retina (Fig. 4A, lane 3) a different pattern was observed, composed of 4 "shadow-like" bands, but this pattern was not clearly seen at other time points in the affected retinas.

Both transcripts were present in testis from a normal and affected dog (Fig. 4A, lanes 10–11, and normal heart (Fig. 4B, lane 10). Very low expression levels were observed in kidney, large intestine, small intestine, stomach and brain, and none was detected in skeletal muscle (Fig. 4B). The 2 transcripts differed by about 600–700 bp, which corresponds to the 655 bp difference between the 2 polyadenylation products identified by 3′ RACE. Thus the 2 bands are presumed to represent polyadenylation variants, and they appear to be co-expressed at similar levels.

3.5. In-silico analysis

The putative canine STK38L protein, like the human, is 464 amino acids long (Supplementary Fig. 1A2, Supplementary Fig. 2). In the N terminus an S100B binding region is identifiable by similarity (Swissprot *in-silico* analysis). This region extends from amino acid 63 to 88. A protein kinase domain extends from amino acid 90 to 383. Three post-translational modification residues are in position 75 (threonine), 282 (serine) and 442 (threonine).

Comparison of the putative canine STK38L sequence to that of 4 other species (human, mouse, chicken and zebrafish) revealed the strong conservation among these species by the high percentage of amino acid identity. Closest to the canine sequence was human (99.1%) followed by mouse (97.4%), zebrafish (90.3%) and chicken (89.8%) (Fig. 5A, Supplementary Fig. 2). In *erd*-affected dogs, 41 amino acids are missing from the putative protein. These 41 amino acids include the complete S100B binding region (26 amino acids, that includes threonine 75), and part of the protein kinase domain (15 amino acids, Fig. 5B).

4. Discussion

In this study, fine mapping followed by positional candidate gene analysis established that *erd* is strongly associated with an *STK38L*/*NDR2* allele characterized by a 4 nt deletion in intron 3 and a SINE insertion in exon 4. The deletion was also found as a minor allele (without the SINE insertion) in canine populations not segregating *erd*. The SINE insertion in exon 4 was only seen in *erd*-affected or -heterozygous dogs — this mutation presumably took place originally on a CFA27 chromosome bearing the 4 nt deletion in intron 3. All *erd*-affected and -heterozygous dogs then inherited this allele identical by descent.

The SINE insertion identified in the *STK38L* gene belongs to the family of SINEC_Cf repeats that comprise a major subclass of canine-

A

	D	*H*	*M*	*Z*	*C*
D		99.1	97.4	90.3	89.8
H	0.9		97.4	90.1	89.4
M	2.6	2.6		89.2	88.3
Z	10.4	10.7	11.7		84
C	11	11.5	12.7	18.1	

B

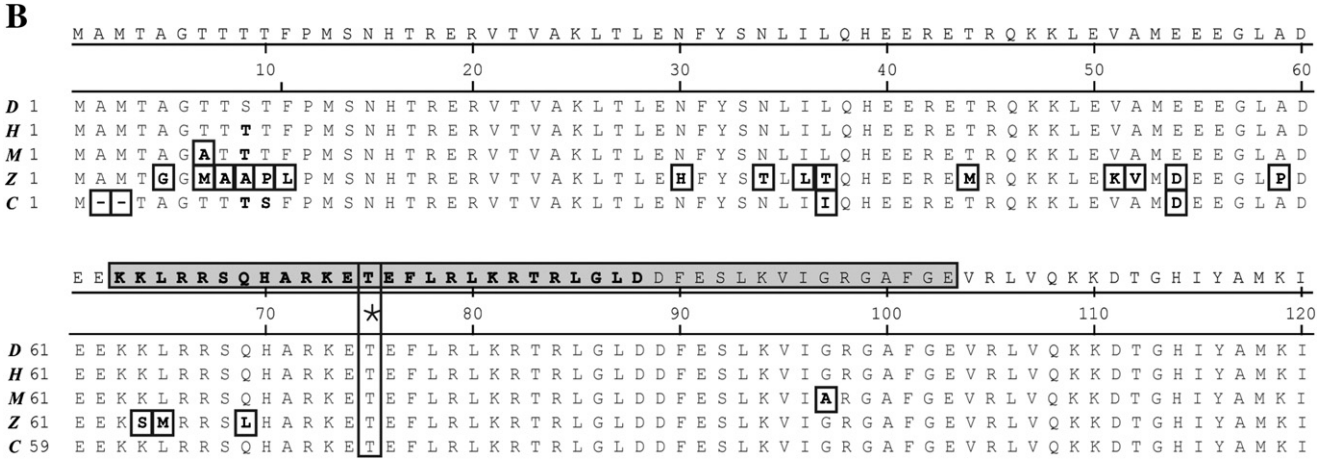

Fig. 5. Conservation analysis of STK38L. A. Pairwise percent nucleotide identities (upper right cells) and divergence (lower left cells) calculated for 5 species (*D* = dog, *H* = human, *M* = mouse, *Z* = zebrafish, *C* = chicken) for the *STK38L* complete coding sequence. The human and canine sequences are the most similar to each other, being 99.1% identical. Sequence distances were calculated using MegAlign 6.1 software (Lasergene, DNAstar, Inc.). B. Multiple sequence alignment of the 5′ end (first 120 codons) of the canine STK38L putative protein sequence (labeled *D*) with that of human (*H*), mouse (*M*), zebrafish (*Z*) and chicken (*C*). The consensus sequence for the 5 species is listed above each set (1...60; 61...120) of aligned sequences. Differences among species are bold and boxed. In the consensus sequence for codons 61–120 (i.e. the lower panel), the 41 amino acids (KKL...FGE) missing from the putative protein in *erd*-affected dogs is boxed and shaded. Within this region, the S100B domain (KKL...GLD) is indicated in bold. The autophosphorylation residue at position 75 is labeled with an asterisk and boxed. The protein kinase domain (FES...) starts at position 90. See Supplementary Fig. 2 for the corresponding analysis of the complete protein. Alignment was done using MegAlign 6.1 software (Lasergene, DNAstar, Inc.).

specific SINEs [12,13]. SINEC_Cf are likely derived from tRNA and contain internal control elements for transcription by RNA polymerase III from a tRNA [14,15]. Nearly half of all annotated genes in CanFam2 contain at least one SINEC_Cf. This element remains so active in the canine genome that of 87,000 young SINEC_Cf elements in CanFam2 nearly 8% are heterozygous [12]. Moreover, comparison of CanFam2 with the previous survey sequence of a standard poodle genome CanFamSS [16] estimate that 20,000 or more of these insertion sites are bimorphic, i.e. not fixed at homozygosity [12,13]. SINE insertions in exonic regions or intron/exon boundaries have previously been implicated in two other canine diseases — narcolepsy [17] and centronuclear myopathy [18] — and in the merle coat color phenotype [19].

The *erd* SINE insertion results in the absence of exon 4 from the mature *STK38L* mRNA, presumably by splicing out during maturation of the mRNA, even though the insertion does not appear to introduce or remove any obvious splicing sites. Furthermore, the mutant transcript appears to be stable in both homozygous and heterozygous retinas, not being degraded by nonsense mediated decay or other mechanisms. The translated protein, however, would clearly be dysfunctional, lacking its S100B binding region, part of its Protein Kinase domain (15 amino acids) and the Thr-75 residue critical for autophosphorylation (Fig. 5B, and Supplementary Figs. 1B2 and 2). Both STK38L/NDR2 and its paralog NDR1 are regulated by binding of S100B, a Ca^{++}-binding protein, to their S100B binding domains [3] and via binding to their N-termini by at least two different human Mob proteins [6,20]. Binding of STK38L by S100B and or Mob proteins is proposed to induce a conformational change that enables STK38L to autophosphorylate at Thr-75, Ser-282, and Thr-444 [5,6]. It is thus reasonable to propose a causative role for this STK38L mutation in *erd*.

Furthermore, the *erd* disease phenotype itself suggests a compatibility between the underlying disease mechanism and proposed roles for the *STK38L/NDR2* pathway in neurons. In the *erd*-affected retina, photoreceptor development is not arrested, but aberrant [1]. Because photoreceptor differentiation occurs postnatally in dogs [21], unlike in humans, the abnormal development of the *erd*-affected retina becomes evident in dogs between 3 and 10 weeks old. The subsequent degeneration of the *erd*-affected retina is initially rapid, but becomes more gradual in dogs over 6 months old [1]. Morphologically, rod and cone inner and outer segments do develop, and their numbers, sizes and proportions on average are initially normal [1]. The primary morphological abnormality during this period of aberrant development is dyscoordination of the length and proportions of inner and outer segments among adjacent rods. This dyscoordination prevents the alignment that normally creates distinct inner and outer segment layers. This alignment starts at around 4 weeks postnatal in normal canine retina, when developing inner and outer segments first become closely packed together, and is essentially complete across the retina by 7–8 weeks postnatal [1,21]. In the same period in *erd*-affected retina, unusual and prominent villiform processes extend from the inner segments of rods and cones, and photoreceptor synapses fail to develop properly, morphologically and functionally. Subsequent to these abnormalities of development, at about the 10th week postnatal, rods and cones degenerate, rapidly at first and later more gradually [1]. The primary pathology in *erd* thus represents dysregulation of photoreceptor growth. Loss of STK38L function makes an appealing explanation for this abnormality.

This is the first time that the STK38L/NDR2 pathway has been implicated in photoreceptor development or disease. With hindsight, however, it is reasonable to expect such a role, given the importance of the polarized, highly regulated, circadian growth and renewal cycle

in photoreceptors, and the proposed roles for STK38L in other, especially neuronal, tissues. STK38L and its interacting partners, S100B and the Mob proteins, thus provide novel candidate genes for hereditary retinal disorders of humans and other species.

Supplementary materials related to this article can be found online at doi:10.1016/j.ygeno.2010.09.003.

Acknowledgments

Technical assistance from Jennifer Johnson, Julie Jordan, Amanda Nickle, and the animal care staff at the RDSF is gratefully acknowledged. Ewen Kirkness (The Institute for Genome Research) kindly provided prepublication information for several canine genomic sequences. This work was supported by NIH Grants EY006855, EY13132, and EY17549; The Foundation Fighting Blindness; the Morris Animal Foundation; and the Van Sloun Fund for Canine Genetic Research.

References

[1] G.M. Acland, G.D. Aguirre, Retinal degenerations in the dog: IV. Early retinal degeneration (erd) in Norwegian elkhounds, Exp. Eye Res. 44 (1987) 491–521.
[2] G.M. Acland, K. Ray, C.S. Mellersh, A.A. Langston, J. Rine, E.A. Ostrander, G.D. Aguirre, A novel retinal degeneration locus identified by linkage and comparative mapping of canine early retinal degeneration, Genomics 59 (1999) 134–142.
[3] R. Tamaskovic, S.J. Bichsel, H. Rogniaux, M.R. Stegert, B.A. Hemmings, Mechanism of Ca2 + -mediated regulation of NDR protein kinase through autophosphorylation and phosphorylation by an upstream kinase, J. Biol. Chem. 278 (2003) 6710–6718.
[4] O. Stork, A. Zhdanov, A. Kudersky, T. Yoshikawa, K. Obata, H.C. Pape, Neuronal functions of the novel serine/threonine kinase Ndr2, J. Biol. Chem. 279 (2004) 45773–45781.
[5] M.R. Stegert, A. Hergovich, R. Tamaskovic, S.J. Bichsel, B.A. Hemmings, Regulation of NDR protein kinase by hydrophobic motif phosphorylation mediated by the mammalian Ste20-like kinase MST3, Mol. Cell. Biol. 25 (2005) 11019–11029.
[6] E. Devroe, H. Erdjument-Bromage, P. Tempst, P.A. Silver, Human Mob proteins regulate the NDR1 and NDR2 serine–threonine kinases, J. Biol. Chem. 279 (2004) 24444–24451.
[7] O. Goldstein, B. Zangerl, S. Pearce-Kelling, D.J. Sidjanin, J.W. Kijas, J. Felix, G.M. Acland, G.D. Aguirre, Linkage disequilibrium mapping in the domestic dog breeds narrows the progressive rod-cone degeneration interval and identifies ancestral disease-transmitting chromosome, Genomics 88 (2006) 541–550.
[8] R. Guyon, T.D. Lorentzen, C. Hitte, L. Kim, E. Cadieu, H.G. Parker, et al., A 1-Mb resolution radiation hybrid map of the canine genome, Proc. Natl Acad. Sci. USA 100 (2003) 5296–5301.
[9] A.V. Kukekova, G.D. Aguirre, G.M. Acland, Cloning and characterization of canine *SHARP1* and its evaluation as a positional candidate for canine early degeneration (*erd*), Gene 312 (2003) 335–343.
[10] T.C. Matise, M. Perlin, A. Chakravarti, Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map, Nat. Genet. 6 (1994) 384–390.
[11] A.V. Kukekova, O. Goldstein, J.L. Johnson, M.A. Richardson, S.E. Pearce-Kelling, A. Swaroop, J.S. Friedman, G.D. Aguirre, G.M. Acland, Canine RD3 mutation establishes rod cone dysplasia type 2 (rcd2) as ortholog of human and murine rd3, Mamm. Genome 20 (2009) 109–123.
[12] K. Lindblad-Toh, C.M. Wade, T.S. Mikkelsen, E.K. Karlsson, D.B. Jaffe, M. Kamal, et al., Genome sequence, comparative analysis and haplotype structure of the domestic dog, Nature 438 (2005) 803–819.
[13] W. Wang, E.F. Kirkness, Short interspersed elements (SINEs) are a major source of canine genomic diversity, Genome Res. 15 (2005) 1798–1808.
[14] M.F. Minnick, L.C. Stillwell, J.M. Heineman, G.L. Stiegler, A highly repetitive DNA sequence possibly unique to canids, Gene 110 (1992) 235–238.
[15] S. Bentolila, J.M. Bach, J.L. Kessler, I. Bordelais, C. Cruaud, J. Weissenbach, J.J. Panthier, Analysis of major repetitive DNA sequences in the dog (*Canis familiaris*) genome, Mamm. Genome 10 (1999) 699–705.
[16] E.F. Kirkness, V. Bafna, A.L. Halpern, S. Levy, K. Remington, D.B. Rusch, A.L. Delcher, M. Pop, W. Wang, C.M. Fraser, J.C. Venter, The dog genome: Survey sequencing and comparative analysis, Science 301 (2003) 1898–1903.
[17] L. Lin, J. Faraco, R. Li, H. Kadotani, W. Rogers, X. Lin, X. Qiu, P.J. de Jong, S. Nishino, E. Mignot, The sleep disorder canine narcolepsy is caused by mutation in the hypocretin (orexin) receptor 2 gene, Cell 98 (1999) 365–376.
[18] M. Pele, L. Tiret, J.-L. Kessler, S. Blot, J.-J. Panthier, SINE exonic insertion in the PTPLA gene leads to multiple splicing defects and segregates with the autosomal recessive centronuclear myopathy in dogs, Hum. Mol. Genet. 14 (2005) 1417–1427.
[19] L.A. Clark, J.M. Wahl, C.A. Rees, K.E. Murphy, Retrotransposon insertion in *SILV* is responsible for merle patterning of the domestic dog, Proc. Natl Acad. Sci. USA 103 (2005) 1376–1381.
[20] S.J. Bichsel, R. Tamaskovic, M.R. Stegert, B.A. Hemmings, Mechanism of activation of NDR (nuclear Dbf2-related) protein kinase by the hMOB1 protein, J. Biol. Chem. 279 (2004) 35228–35235.
[21] G.D. Aguirre, L.F. Rubin, S.I. Bistner, Development of the canine eye, Amer. J. Vet. Res. 33 (1972) 2399.