2017

A 16.7 kb deletion in *Sipa1l3* is associated with juvenile cataract in mice

Lianna R. Walker
University of Nebraska–Lincoln, lwalker8@unl.edu

Emily R. Tosky
University of Nebraska–Lincoln

Kylee M. Sutton
University of Nebraska–Lincoln

Rhonda Griess
University of Nebraska–Lincoln, rgriess3@unl.edu

Marytza D. Abebe
University of Nebraska–Lincoln

See next page for additional authors

Published in *Mammalian Genome* 28 (2017), pp. 515–519. doi 10.1007/s00335-017-9720-9

Submitted 17 August 2017; accepted 21 September 2017; published online 26 September 2017.

PMID: 28951961

A 16.7 kb deletion in *Sipa1l3* is associated with juvenile cataract in mice

Lianna R. Walker,[1] Emily R. Tosky,[1] Kylee M. Sutton,[1] Rhonda Griess,[1] Marytza D. Abebe,[1] Sarah Y. Barnes,[1] Tom Cunnigham,[2] Stephen D. Kachman,[3] Merlyn K. Nielsen,[1] and Daniel C. Ciobanu[1]

1 Animal Science Department, University of Nebraska–Lincoln
2 Molecular Resource Center, University of Tennessee Health Science Center, Memphis, TN 38163
3 Department of Statistics, University of Nebraska–Lincoln

Corresponding author — Daniel C. Ciobanu; dciobanu2@unl.edu

Abstract

Congenital or juvenile cataract is a disease condition in which opacification of the lenses is present at birth or manifests early in life. It has been attributed to different monogenic factors with a high degree of heterogeneity and is often studied using mouse models. A spontaneous mutation was identified in a mouse line selected for heat loss that influenced lens formation and resulted in juvenile cataracts in mice homozygous for the recessive allele. Genetic dissection of this selection line by combining high-density genotypes and homozygosity mapping uncovered a 906 kb fragment on MMU7 encompassing 21 SNPs split into two groups of consecutive, homozygous segments specific to the cataract phenotype. Haplotype analysis revealed a 197.5 kb segment unique to cataract-affected mice that included a single known transcript consisting of the first 14 exons of *Sipa1l3*. In this region, we discovered a deletion of 1114 bp at the mRNA level, spanning four coding exons, predicted to produce a truncated *Sipa1l3* protein lacking a portion of a Rap-GAP domain and two other potentially vital domains. At the genome level, the deletion consisted of 16,733 bp. Genotyping across different samples confirmed that only affected mice were homozygous for the deletion and normal mice were either heterozygous or homozygous for the wild-type allele. Further studies will be required to determine the impact of the truncated *Sipa1l3* domains on eye development.

Background

Cataract is a disease condition characterized by the opacification of the lenses. Congenital or juvenile cataract occurs at birth or early in life with a frequency of one to six cases occurring per 100,000 births. A large proportion of congenital or juvenile cataract conditions have a genetic source that is usually monogenic, but also characterized by heterogeneity (Pichi et al. 2016). Currently, there are more than 20 known genes involved in congenital, infantile or juvenile cataract (Evers et al. 2015). Mouse models have been used extensively in genetic research of these conditions, with over 400 reported loci associated with cataract across different strains including spontaneous and chemically or radiation-induced mutations (http://www.informatics.jax.org/). A spontaneous recessive mutation influencing lens formation was identified in a mouse line selected for heat loss (Cargill et al. 2001) with an original genetic base comprising of four outbred populations (Jones et al. 1992; Nielsen et al. 1997). Mice homozygous for the recessive allele exhibited complete cataracts at a juvenile age (Fig. 1). This mutation, called *jrc* for juvenile recessive cataract mutation, was found to be linked to a tyrosinase (*Tyr*) recessive allele located on MMU7 that caused albino coat color. The position of the *jrc* locus was originally mapped between *D7Mit340* (4.63 Mb) and *D7Mit227* (36.58 Mb) (Cargill et al. 2001) with a follow-up by Poskochil (2004, Undergraduate Honors Thesis, personal communication) who suggested that *D7Mit266* (28.3 Mb) was the closest marker to *jrc*. An unknown protein was found to be absent in both water-soluble and insoluble protein profiles of the mice with cataracts compared to normal mice (Tebbe 1999, Undergraduate Honors Thesis, personal communication). The eyes of the *jrc* mice were smaller and the lenses were more fragile, suggesting that the missing protein may be a component of the fiber cells that lead to opacity and fragility (Tebbe 1999). The objective of this research was to combine high-density genotypes generated by MegaMUGA Universal Genotyping Array with homozygosity mapping in order to dissect the *jrc* locus and identify the gene and causal variant associated with this type of juvenile cataract in mice.

Materials and methods

The experimental design and procedures used during this research project were approved by the Institutional Animal Care and Use Committee of the University of Nebraska–Lincoln (UNL).

Population resource, phenotype, and tissue collection

The samples and phenotypes collected in this study originated from an albino line with cataract (BD) and from the offspring of the 51st generation of a cross (BY) between BD and the C57BL/6J from University of Nebraska–Lincoln. The albino line was derived from a mouse line selected for heat loss and originating from a four-way cross of outbred populations (NIH,

ICR, CF-1, and CFW) (Cargill et al. 2001). Ten breeding cages, with two females and one male assigned per cage, were used to produce subsequent generations with limited inbreeding. The presence of cataract was recorded at 3 weeks of age. Tail samples were collected for genotyping purposes. Lenses were collected in RNAlater (Thermo Fisher Inc.) from selected 3-week-old individuals that displayed normal or cataract eyes.

Genotyping and homozygosity mapping

Genomic DNA was isolated from tail samples using the DNeasy kit (Qiagen). DNA quality was assessed using an Epoch Microplate Spectrophotometer (BioTek). High-density genotyping of juvenile cataract (n = 10) and normal (n = 6) mice was performed by MegaMUGA Bead Array (Illumina Inc.) that included 77,300 SNPs. The genotyped mice represented eight different litters and both sexes. Six of the litters include mice with both cataract and normal phenotypes. Additional MegaMUGA genotypes from a variety of laboratory mouse lines exhibiting normal phenotypes were downloaded from UNC System Genetics database (http://csbio.unc.edu/CCstatus/index.py?run=Geneseek). Monomorphic SNPs and SNPs with a call rate < 80% were excluded leaving 27,222 SNPs for further analysis.

Homozygosity mapping was conducted in JMP (SAS Inc.) in order to identify common homozygous segments in cataract-affected mice. Haplotype analysis of cataract-affected and normal mice, including five common laboratory inbred strains (A/J, BALB/cByJ, C3H/HeJ, C57BL/6J, and DBA/2J), was carried out using Haploview 4.2 (Barrett et al. 2005) to determine haplotypes in the region associated with the largest number of cataract-specific homozygous SNPs. Potential candidate genes and their ontologies in the targeted region were obtained using the Ensembl BioMart data-mining tool (http://uswest.ensembl.org/biomart/martview/9883b5f936490aef948937cd5de40e1b).

cDNA sequencing

Total RNA was extracted from pooled lenses using TRIzol (Thermo Fisher Scientific Inc.). Complementary DNA (cDNA) was obtained using First strand cDNA Synthesis (GE Healthcare). Primer 3 software and a *Sipa1l3* reference mRNA sequence (NM_001081028.1) were used for primer design. Amplification of cDNA samples (GoTaq Flexi DNA polymerase, Promega; Forward 5′ GTA CAA CGA GGA AGC CG 3′, Reverse 5′ TGG TTT CCG ACT CCG TCT TG 3′) across *Sipa1l3* in both normal and affected samples was evaluated by agarose electrophoresis. PCR products were subjected to ExoSAP-IT (USB Corporation) and sequenced using dye terminators and ABI PRISM 3130 Genetic Analyzer (Applied Biosystems). Sequencher software (Gene Codes) was used for aligning cDNA sequences and polymorphism identification.

Gene expression profiling

Expression of *Sipa1l3* was evaluated by real-time quantitative PCR (qPCR) using the cyclophilin D (*Ppid*) gene for normalization. Mean normalized expression (MNE) values were obtained based on cycle crossing thresholds (Cq) from technical triplicates taking the efficiency of qPCR into account (Simon et al. 2002). The MNE values of normal (n = 4) and cataract (n = 4) samples were log10 transformed and compared using a t test.

Validation assay

A three primer PCR was used to genotype the *Sipa1l3* indel in 9 normal (BY, C57B6/J, and DBA/2J) and 16 cataract-affected (BY and BD) mice by agarose gel electrophoresis. Primer 3 software, the reference *Sipa1l3* sequence (NM_001081028.1), and sequences obtained from cataract- affected mice were used for primer design (Forward 5′ GTA CAA CAA CGA GGA AGC CG 3′, Reverse-1 5′ AGC AGG TCA GTC AGT CAA GC 3′, Reverse-2 5′ TCT GCC CTA TCC ACT GCT CT 3′). Genomic DNA was isolated from tail samples using the DNeasy kit (Qiagen).

Results

Homozygosity mapping

In order to identify the genomic region responsible for the juvenile cataracts, we used an identity by descent (IBD) mapping approach using high-density genotyping generated by MegaMUGA Bead Array of normal and mice displaying juvenile cataract. Out of 27,222 informative SNPs, there were 21 SNPs where all the affected mice were homozygous and the normal mice were heterozygous or homozygous for the alternate allele. All the SNPs that fit this stringent criterion in this highly related population were located on a 906 kb fragment on MMU7 between 29,589,840 (rs48074914, mm9) and 30,496,080 bp (rs36304650). With the exception of the last SNP (rs36304650), the rest of these SNPs clustered in two groups of 12 and 8 SNPs, respectively. The largest segment includes 12 consecutive SNPs spanning 299.3 kb between 30,034,098 (rs31882446) and 30,333,369 bp (rs50788430). This segment includes 4 genes, *Spint2, Ppp1r14a, Dpf1*, and *Sipa1l3*. Using the GeneAtlas MOE430 mouse microarray expression data set (Lattin et al. 2008), we found that the only gene that had substantial expression in the lenses of C57BL/6J mice was Signal Induced Proliferation Associated 1 Like 3 *(Sipa1l3)*. The expression of *Sipa1l3* in the lenses was the highest of all 191 tissues profiled, 13.7 fold greater than the tissues median.

Haplotype analysis was used to fine map the region associated with juvenile cataract, including all BD/BY cataract and normal mice as well as 5 inbred strains with a normal phenotype that are used extensively as parents of various genetic resource populations, including A/J (Collaborative

Cross, AXB/BXA, BXH), BALB/cByJ (CXB), C3H/HeJ (BXH), C57BL/6J (Collaborative Cross, AXB/BXA, BXD), and DBA/2J (BXD). Haplotype analysis of the region covered by the 21 SNPs with a specific homozygous status in the affected mice showed that a 197.5 kb region covered by 11 SNPs located between 30,135,854 and 30,333,369 bp was the only segment with a haplotype specific to cataract-affected mice. The only known transcript located in this region included the first 14 exons of *Sipa1l3*. Eight of these MegaMUGA SNPs were located in the introns of *Sipa1l3*, and with only one exception all affected mice had a specific homozygote genotype.

cDNA and genomic DNA sequencing

Sipa1l3 is one of the largest genes in the mouse genome (185,083 bp), including 27 exons, one known mRNA isoform (NM_001081028.1; 7713 bp) with a predicted peptide of 1776 amino acids. Sequencing the cDNA isolated from lenses in normal and affected mice uncovered 4 SNPs, a 3 bp indel, and a large deletion across the coding exons of *Sipa1l3*. The SNPs included 1 missense (rs46083377) and three synonymous (rs31232119, rs49407811, rs32357946) polymorphisms. The genotypes of these SNPs and the 3 bp indel were shared between affected and normal phenotypes. Analysis of known polymorphisms located in the coding region of *Sipa1l3* using Sanger's Mouse Genome Project (http://www.sanger.ac.uk/science/data/mouse-genomes-project) showed that most of these polymorphisms were already identified using sequencing data derived from 18 common mouse laboratory strains, all with normal eyes.

The large deletion covered exons 8–11, spanning 1114 bp of the *Sipa1l3* mRNA (Fig. 2). The truncated allele (c2628_3741del,p.Thr671_Leu1776delinsArgGlyGlyGlnArgProMetIle) is predicted to encode a 678 amino acid peptide compared to the 1776 amino acid peptide specific to the wild-type allele. At a genome level, the length of the deletion is 16,733 bp (GRCm38/mm10-chr7:29,370,216-29,386,948; Fig. 2). Genotyping the 16.7 kb indel in a subset of samples with normal and cataract eyes showed that only the cataract-affected mice were homozygous for the truncated *Sipa1l3* allele (Fig. 3).

Quantitative analysis of gene expression

Expression profiling of *Sipa1l3* in lenses revealed that the truncated *Sipa1l3* allele resulted in nonsense-mediated mRNA decay since a reduction in expression by 88.6% was observed in cataract-affected samples compared to mice with normal phenotypes (P = 0.0009; Fig. 4). While in this study we did not quantify the expression of the truncated *Sipal13* peptide, an early report found an unknown protein absent in both water-soluble and insoluble protein profiles of the mice with cataracts compared to normal mice (Tebbe 1999). We hypothesize that *Sipal13* may be the missing protein.

Discussion

Sipa1l3 has a role in epithelial cell morphogenesis, establishing or maintaining cell polarity, cell adhesion, and organization of the cytoskeleton in lenses. The mouse protein encoded by *Sipa1l3* is predicted to include a Rap GTPase-activating protein (Rap-GAP) domain, a PDZ domain, and a C-terminal domain of the SPAR protein (Greenlees et al. 2015). The Rap-GAP domain is specific for Rap1, which is a GTP-binding protein and suggested, like other small GTPases, to be involved in functions associated with lens development (Rao et al. 1997). The PDZ domain was predicted to interact with 14-3-3ζ (YWHAZ), an isoform from a family of essential intracellular signaling molecules in human lenses (Evers et al. 2015) and suggested to be involved in cataract formation (Lichtstein et al. 2000). The role of the C-terminal domain of the SPAR protein is less known. The truncated *jrc Sipa1l3* allele includes only the first 35 bp of the 180 bp Rap-GAP domain and completely lacked the PDZ domain and the C-terminal domain of the SPAR protein.

The peptide sequence encoded by *Sipa1l3* is conserved across mammalian species with the same domains predicted in the human ortholog. A recent study employing exome sequencing uncovered a nonsense variant (c.4489C>T; p.R1497*) in *Sipa1l3* in siblings with congenital cataract (Evers et al. 2015). The peptide encoded by this variant is truncated and contains the Rap-GAP and PDZ domains, but includes only a small fragment of the C-terminal domain of the SPAR protein. Recently, Greenlees et al. (2015) found that dysfunction in *Sipa1l3* in human, zebrafish, and mouse, affected lens and eye development. For example, a missense mutation (p.Asp148Tyr) and a balanced chromosomal translocation affecting the 5′UTR of *Sipa1l3* were found in patients with congenital cataract and other eye abnormalities. The missense mutation affected a putative actin-binding site while the chromosomal translocation resulted in reduced *Sipa1l3* expression (Greenlees et al. 2015).

At first glance the *jrc* variant appears to be more severe, leading to partial or complete truncation of all conserved *Sipa1l3* domains. It is less clear if the early cataract is a result of the truncated protein being abnormally folded or if the missing domains influence interactions with other critical proteins such as Rap1 or 14-3-3ζ. Evers et al. (2015) found that a truncated *Sipa1l3*, lacking the last 249 amino acids of the C-terminal domain of the SPAR protein, was sufficient to cause cataracts. Future studies will need to establish the specific impact of each domain in eye development and pathways that lead to early cataract formation.

Conclusion

By incorporating high-density genotypes in both homozygosity mapping and haplotype analyses, we identified a novel 16.7 kb deletion on MMU7 associated with juvenile cataract in a mouse line selected for heat loss. This deletion spans 4 coding exons of *Sipa1l3*, one of the largest genes in the

mouse genome believed to play a critical role in lens development. The deletion is predicted to result in a truncated *Sipa1l3* protein lacking functional domains highly conserved across species. Other genetic variants leading to dysfunction of the *Sipa1l3* protein have been implicated in cataract formation and abnormal eye development in various species, including human. Further studies will be required to determine the exact manner in which this or other deletions disrupt important domains and proper *Sipa1l3* function leading to cataract.

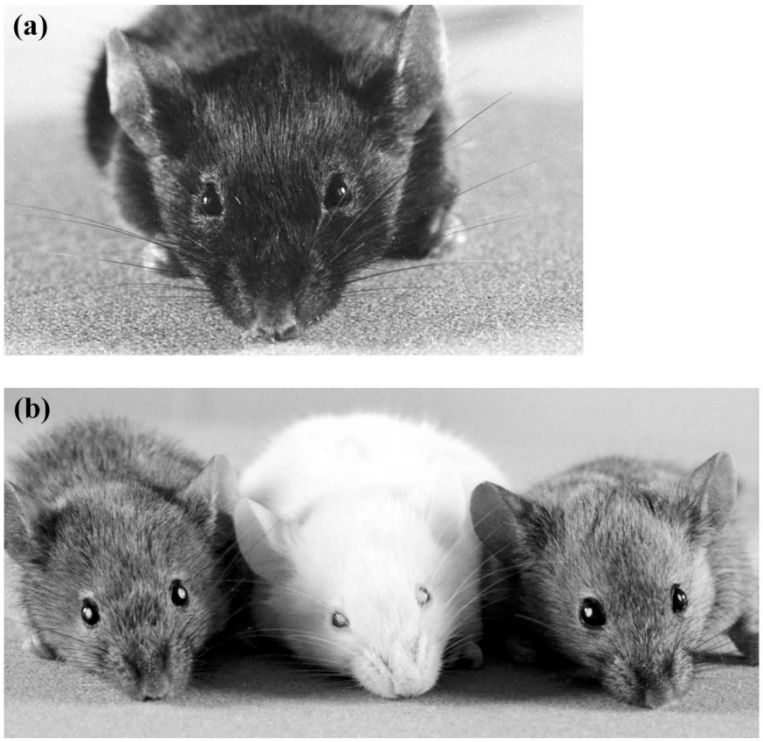

Fig. 1 Mice from BY line displaying **(a)** normal eyes and **(b)** juvenile cataract.

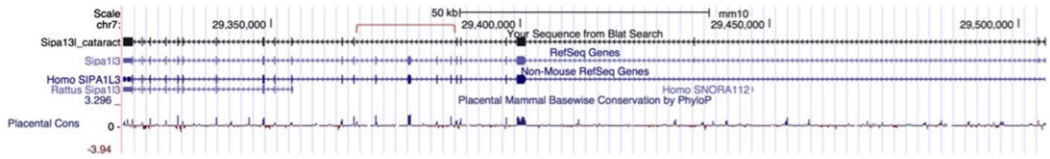

Fig. 2 Gene structure of *Sipa1l3* in the wild-type and cataract variants. Cataract variant is missing exons 8–11 compared to wild-type sequence (red bracket).

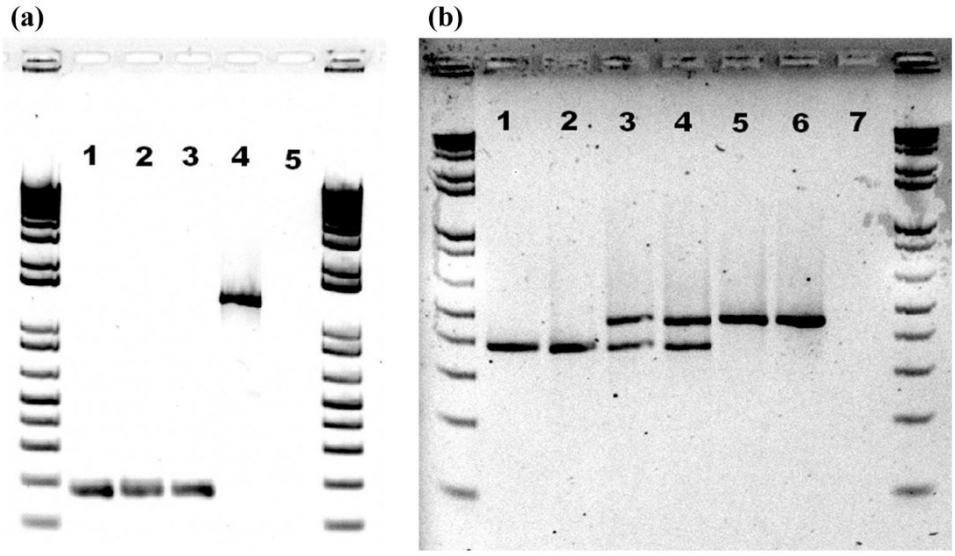

Fig. 3 Electrophoresis profiles of the **(a)** cDNA in cataract-affected (lanes 1–3) and normal (lane 4) genotypes and **(b)** genomic DNA in cataract-affected (lanes 1–2) and normal (lanes 3–6) genotypes. Negative controls are located in lane 5 (**a**) and lane 7 (**b**). A 1 kb plus DNA ladder (Invitrogen) is represented in the first and last lanes of both images.

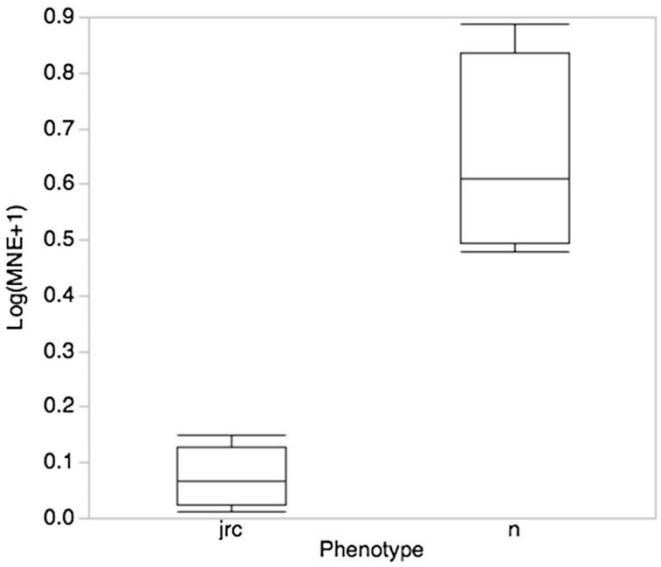

Fig. 4 *Sipa1l3* gene expression in normal (n) and cataract-affected mice (*jrc*). Log10 transformed mean normalized expression (MNE) of *Sipa1l3* mRNA extracted from mice lenses is represented on the y-axis with corresponding phenotype represented on the *x*-axis.

Acknowledgments — Financial support for this research was provided by Undergraduate Creative Activities and Research Experiences (UCARE) Program of the University of Nebraska–Lincoln, the Nebraska Tobacco Settlement Biomedical Research Development Funds and University of Nebraska–Lincoln Agricultural Research Division. All sequencing data were generated in the Molecular Resource Center of Excellence at The University of Tennessee Health Science Center.

References

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265. doi:10.1093/bioinformatics/bth457

Cargill EJ, Happold TR, Bertani G, Rocha J, Lou MF, Pomp D, Nielsen MK (2001) Localization of a recessive juvenile cataract mutation to proximal chromosome 7 in mice. Hum Hered 52:77–82

Evers C et al (2015) SIPA1L3 identified by linkage analysis and whole-exome sequencing as a novel gene for autosomal recessive congenital cataract. Eur J Hum Genet 23:1627–1633. doi:10.1038/ejhg.2015.46

Greenlees R et al (2015) Mutations in SIPA1L3 cause eye defects through disruption of cell polarity and cytoskeleton organization. Hum Mol Genet 24:5789–5804. doi:10.1093/hmg/ddv298

Jones LD, Nielsen MK, Britton RA (1992) Genetic variation in liver mass, body mass, and liver:body mass in mice. J Anim Sci 70:2999–3006

Lattin JE et al (2008) Expression analysis of G Protein-Coupled Receptors in mouse macrophages. Immunome Res 4:5. doi:10.1186/1745-7580-4-5

Lichtstein D, McGowan MH, Russell P, Carper DA (2000) Digitalis and digitalis-like compounds down-regulate gene expression of the intracellular signaling protein 14-3-3 in rat lens. Hypertens Res 23:S51–S53

Nielsen MK, Jones LD, Freking BA, DeShazer JA (1997) Divergent selection for heat loss in mice: I. Selection applied and direct response through fifteen generations. J Anim Sci 75:1461–1468

Pichi F, Lembo A, Serafino M, Nucci P (2016) Genetics of congenital cataract. In: Pediatric cataract, vol 57. Karger, Basel, pp 1–14

Rao PV, Zigler JS Jr, Garland D (1997) Analysis of small GTP-binding proteins of the lens by GTP overlay assay reveals the presence of unique GTP-binding proteins associated with fiber cells. Exp Eye Res 64:219–227. doi:10.1006/exer.1996.0197

Simon MF, Rey A, Castan-Laurel I, Gres S, Sibrac D, Valet P, Saulnier-Blache JS (2002) Expression of ectolipid phosphate phosphohydrolases in 3T3F442A preadipocytes and adipocytes. Involvement in the control of lysophosphatidic acid production. J Biol Chem 277:23131–23136. doi:10.1074/jbc.M201530200