

Identification of a *MIP* mutation that activates a cryptic acceptor splice site in the 3' untranslated region

Chongfei Jin,¹ Jin Jiang,² Wei Wang,¹ Ke Yao¹

¹Eye Center of the Second Affiliated Hospital, Medical College of Zhejiang University, Hangzhou, China; ²Department of Ophthalmology, Zhejiang Provincial People's Hospital, Hangzhou, China

Purpose: To investigate the consequence of a major intrinsic protein *MIP* splice-site mutation (c.607-1G>A) in a four-generation Chinese pedigree afflicted with autosomal dominant congenital cataracts (ADCC).

Methods: Both a mutated minigene with c.607-1G>A, and a wild-type minigene were constructed using the pTARGET™ mammalian expression vector. They were transiently transfected into HeLa cells and human lens epithelial cells, respectively. After 48 h incubation, RNA extraction and RT-PCR analysis were performed and PCR products were separated and confirmed by sequencing. Structural models of the wild-type and the mutant aquaporin 0 (AQP0) were generated and analyzed using SWISS-MODEL.

Results: The G>A transition activated a cryptic acceptor splice site (c.965-966) in the 3' untranslated region (3' UTR), resulting in the absence of the coding region and most of the 3'UTR in exon 4 of the mature mRNA. Moreover, homology modeling of the mutant protein suggested that the sixth transmembrane helix and carboxyl terminus were replaced with the Leu-His-Ser tripeptide (AQP0-LHS).

Conclusions: The *MIP* splice-site mutation (c.607-1G>A) activates a cryptic acceptor splice site in the 3' UTR, which may result in substitution of the sixth transmembrane helix and carboxyl terminus for AQP0-LHS. To our knowledge, this is the first report of activation of a cryptic splice site in the 3' UTR in a human disease gene.

Hereditary congenital cataract (OMIM 604307) is an opacification of the eye lens which presents at birth or shortly thereafter. Despite the great advances in understanding of lens structure and function, the relationships among cataract morphology, etiology, and pathologic mechanisms are still unclear. Congenital cataracts remain the leading cause of visual disability in children worldwide [1-4]. In 2000, the major intrinsic protein (*MIP*, MP26, AQP0) was reported to be associated with autosomal dominant congenital cataract (ADCC) [5]. To date, seven mutations in the *MIP* gene have been identified from seven unrelated human families; five missense mutations (T138R, E134G, R33C, R233K, and V107I) [5-8], one deletion mutation causing a frameshift that alters 41 of 45 subsequent amino acids and creates a premature stop codon (G213VfsX45) [9], and one splice-site mutation causing autosomal dominant congenital cataract (ADCC) with "snail-like" phenotype in a large Chinese family (our laboratory's recently reported finding, c.607-1G>A) [10].

The effect of mutations that alter critical amino acids and generate premature stop codons can be predicted from triplet codon analysis. However, it is difficult to predict the effects of changes in intronic sequences. The production of mature mRNA requires an accurate and efficient removal of introns

from pre-mRNA by the spliceosome [11]. The regulation of alternative splicing usually relies on a choice from among candidate splice sites, either alternative 5' or alternative 3' sites. Mutations can disrupt splicing by directly inactivating or creating an alternative splicing site, by activating a cryptic splice site or by interfering with splicing regulatory elements. Although the activation of cryptic splice sites has been reported as a mechanism in human disease, to the best of our knowledge currently there is no report of activation of cryptic splice site in the 3' untranslated region (3' UTR, sequences on the 3' end of mRNA which are not translated into protein) [12].

METHODS

Plasmid and minigene constructs: A mutated minigene was obtained by the insertion of a PCR fragment containing the genomic *MIP* sequence from the proband into the pTARGET™ mammalian expression vector (Promega, Milan, Italy). It consisted of the 5' flanking sequence of exon 3, exon 3, intron 3, and exon 4 including the coding and the 3-untranslated region (3'UTR) which was amplified using primers MIPS (5'-TTG ACC CCA AGG TAG AAA TGA CG-3') and MIPAS (5'-ACA ATC AGC ACA CAC GGC GAA G-3'). PCR was performed on 100 µg genomic DNA in a standard 50 µl volume, containing 1.25 PrimeSTAR® HS DNA Polymerase (Takara Bio, Otsu, Japan). The cycling conditions for PCR were 35 cycles of 98 °C for 10 s, 68 °C for 4 min, preceded by 30 s at 98 °C and followed by a final elongation step at 68 °C for 10 min. The PCR product was

Correspondence to: Ke Yao, Ph.D., M.D., Eye Center of the Second Affiliated Hospital, Medical College of Zhejiang University, 88 Jiefang Road, Hangzhou, China, 310009; Phone: +86-571-87783897; FAX: +86-571-87783908; email: xlren@zju.edu.cn

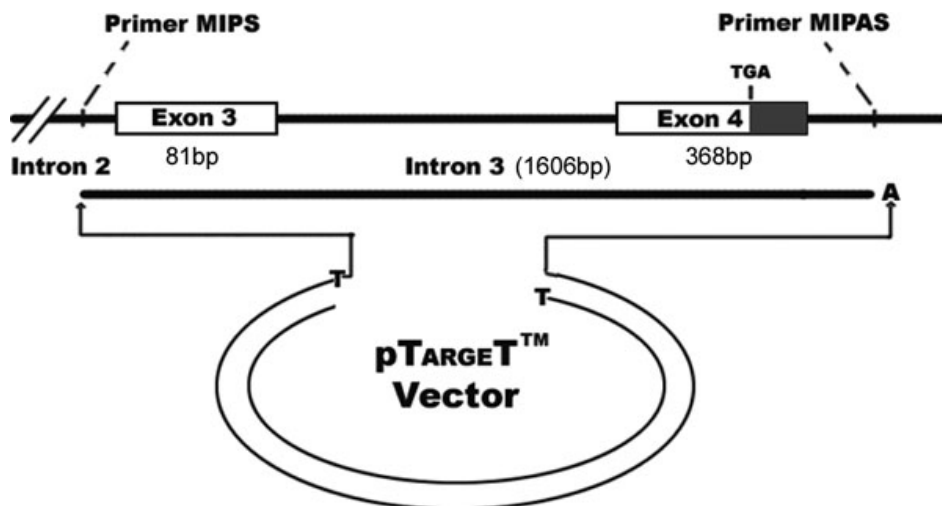


Figure 1. Schematic diagram of the minigene construct. It was obtained by insertion of a PCR fragment containing the genomic *MIP* sequence into the pTARGET mammalian expression vector after the A-tail was added. It consists of the 5' flanking sequence of exon 3, exon 3, intron 3, and exon 4 including the coding and the 3'-untranslated region (3'UTR) which was amplified using primers MIPS and MIPAS.

cloned into the vector after the addition of an A-tail (Figure 1). The addition of the A-tail was achieved by incubating 5 μ l purified PCR product in a 10 μ l volume, containing dATP (final concentration 0.2 mM) and 5 U Taq DNA Polymerase (Takara Bio). The temperature condition was at 70 °C for 30 min. Plasmid DNA was isolated using a Miniprep Kit (Axygene, Union City, CA) and *MIP* c.607-1G>A was ascertained by sequencing—pTARGET-mut. A similar approach was used to obtain the corresponding wild type minigene by amplifying the genomic DNA from a healthy individual—pTARGET-wt. Each clone was entirely sequenced to confirm that no other mutations had been introduced by the PCR procedure.

Cell culture and transient transfection: HeLa cells and human lens epithelial cells (LECs, HLE B-3) were cultured in RPMI 1640 supplemented with 10% fetal bovine serum. Cells were grown in a humidified atmosphere of 5% CO₂ and 95% air at 37 °C and cultured according to standard procedures. Transient transfections were performed in 10-cm dishes (2 \times 10⁶ cells/dish) with Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA) using 4 μ g of total plasmid DNA Endofree purified (Qiagen, Hilden, Germany) following the manufacturer's instructions.

Minigene splicing assay: Both HeLa cell and LECs were transfected with pTARGET-mut and pTARGET-wt minigene constructs. The blank minigene construct pTARGET-con was used as a control. Forty-eight hours after transfection, total RNA from the cells was extracted with an RNeasy Mini Kit (Qiagen) and analyzed by RT-PCR. Reverse transcription was performed using M-MuLV reverse transcriptase (Fermentas, Vilnius, Lithuania) and oligo (dT) primer in a 20 μ l volume according to the protocol of the manufacturer, and the PCR reaction was performed with a forward primer (5'-CTT TGC TCC TGC CAT TCT-3') and a vector-specific reverse primer (5'-GGC TTT ACA CTT TAT GCT TC-3').

PCR products were separated by electrophoresis on a 2% agarose gels and confirmed by sequencing.

Comparative modeling of AQP0: The three-dimensional structure of AQP0 was modeled on the basis of the crystal structure of bovine AQP0. The homology model was generated by SWISS-MODEL and analyzed in the Swiss-PdbViewer, Version 3.7 [13-15].

RESULTS

Effect of the intronic mutation on the processing of the pre-mRNA transcript: Figure 2 showed the results when plasmids, pTARGET-mut, pTARGET-wt, and pTARGET-con were independently transfected to HeLa cells and LECs (not expressing AQP0), and total RNA was extracted from transfected cells and analyzed by RT-PCR followed by electrophoresis on 2% agarose gels. RT-PCR of the total RNA obtained from cells transfected by pTARGET-wt yielded a 687 bp band consistent with the correct splicing of pre-mRNA. This was confirmed by sequencing of the isolated PCR product (Figure 3A). In contrast, a 327 bp band was detected in cells transfected with pTARGET-mut (Figure 2) and the sequencing analysis shown in Figure 3B demonstrated that a cryptic acceptor splice site was activated at nt359 of exon 4 (c.965-966) in the 3'UTR of the *MIP* gene. A summary of these results was shown in Figure 4.

Comparison of wild-type and mutant AQP0 structural models: A mutant protein with 205 amino acids was supposed to be generated when the cryptic splice site in the 3' UTR was activated. Wild-type AQP0 structural model contains six transmembrane α -helical domains with a short 9 amino acid NH₂-terminus and a longer 39 amino acid COOH-terminus extending into the cytoplasm, whereas the mutant AQP0 structural model showed the loss of the last α -helical domain as well as the complete COOH-terminus. Instead, the aberrant AQP0 model (AQP0-LHS) was ended up having the Leu-His-Ser tripeptide (Figure 5).

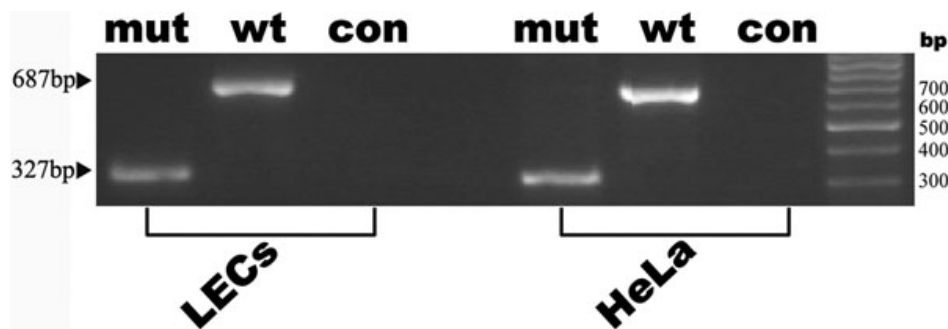


Figure 2. RT-PCR products obtained from minigene constructs transient transfection. The mutated (pTARGET-mut) and wild-type (pTARGET-wt) minigene constructs were independently transfected into HeLa cells and human lens epithelial cells (LECs). Forty-eight hours later, total RNA was extracted from transfected cells and analyzed by RT-PCR. The RT-PCR product obtained from pTARGET-wt and pTARGET-mut shows a 687 bp band and a 327 bp band, respectively. The pTARGET-col used as a control shows a negative band.

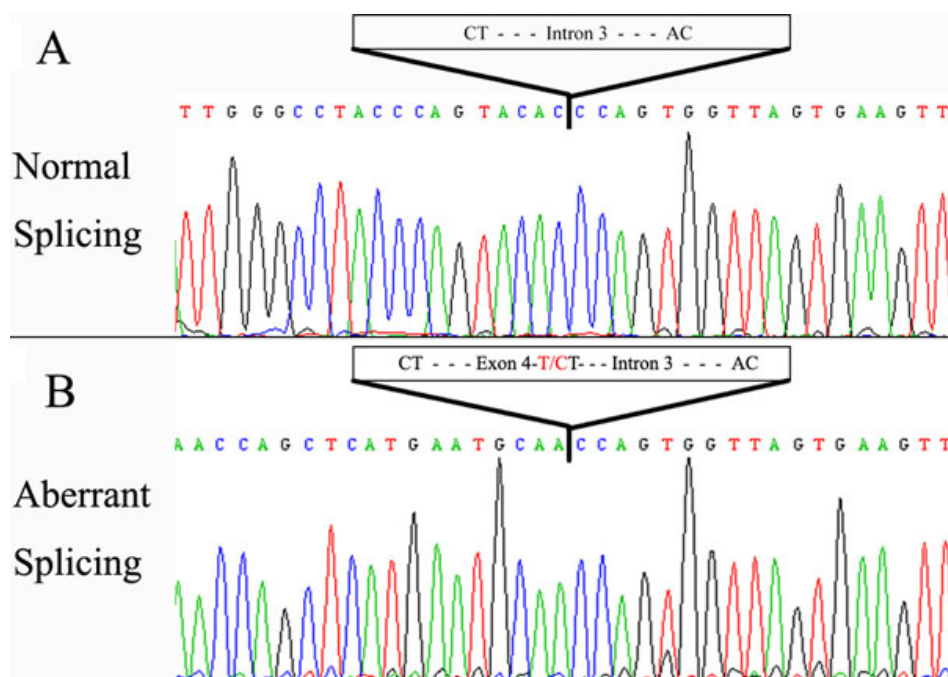


Figure 3. Reverse sequence of the wild-type and the mutant RT-PCR products. **A:** Normal splicing shows the exact excision of intron 3. **B:** Aberrant splicing shows the activation of a cryptic splice acceptor site within the 3' UTR of the *MIP* gene, resulting in the loss of the coding region and most of the untranslated region of exon 4.

DISCUSSION

It is reported that mutations in acceptor splice sites can result in exon skipping or activation of cryptic sites, either upstream or downstream of authentic sites [16]. In the present study, there was a mutation in the invariant AG dinucleotide of the last intron, so none of the existing canonical AG sites downstream are available at the 5' border of following exon as an alternative acceptor splice site. According to previous studies, it is advisable to identify possible splicing defects by analyzing the mRNA available from the individual patients [17]. Therefore, we collected blood specimens and the cataracts from the proband during surgery. However, we failed to extract *MIP* mRNA from the blood specimens or the cataract probably due to the lower mRNA expression levels in the mutant alleles (unpublished). To confirm the role of this

mutation, we finally made use of a simple minigene splicing assay [18]. Subsequently, we demonstrated that the acceptor splice-site mutation (c.607-1G>A) resulted in activation of a cryptic splice site in the 3' UTR of the *MIP* gene. To our knowledge, cryptic acceptor splice site in the 3' UTR has not been reported in human disease genes to date (summary in Table 1) [12,16]. According to [Human Splicing Finder](#) [19], the natural acceptor splice site is scored 96.46, while the cryptic acceptor splice site is scored 74.48. Interestingly, there is a consensus motif (TCTTTC) upstream of the acceptor splice site. Further study is required to determine whether there is a role for the TCTTTC motif in the aberrant splicing.

The activation of a cryptic acceptor splice site may lead to substitution of the distal end of the 6th trans-membrane domain and the COOH-terminal domain of AQP0 for a Leu-His-Ser tripeptide (AQP0-LHS). Previous studies have

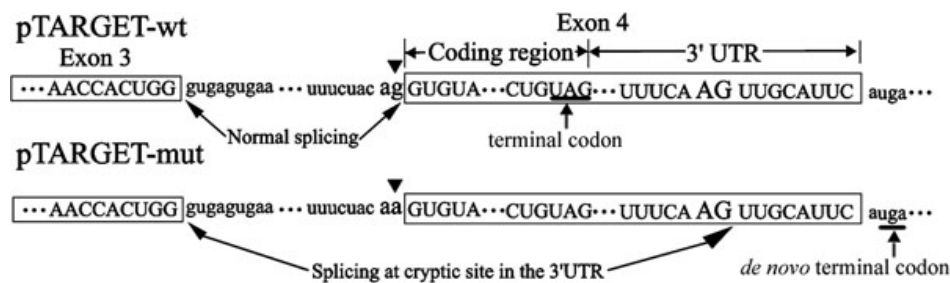


Figure 4. Schematic diagram of comparison between normal splicing and aberrant splicing. The upper panel shows the normal splicing of the wild-type RNA transcript. The lower panel shows that the G>A mutation (black triangles) at the acceptor splice site of intron 3 leads to activation of a cryptic acceptor splice site within the 3' UTR of the *MIP* gene, resulting in the loss of the coding region and most of the untranslated region of exon 4. The underlining codon (UAG and uga) denote the wild type and de novo terminal codon, respectively. The 61 amino acids encoded by exon 4 are replaced by a Leu-His-Ser tripeptide.

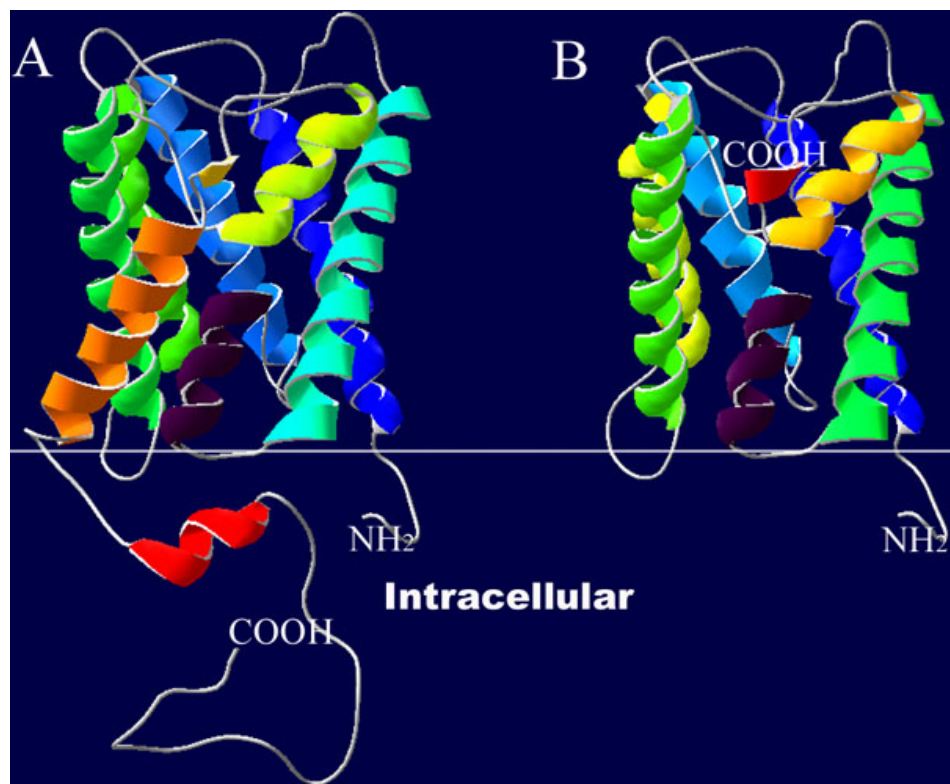


Figure 5. Comparison of wild-type and mutant AQP0 structural models by SWISS-MODEL. **A:** The wild-type AQP0 structural model contains six transmembrane α -helical domains with a short 9 amino acid NH₂-terminus and a longer 39 amino acid COOH-terminus extending into the cytoplasm. **B:** The mutant AQP0 structural model shows the absence of the last α -helical domain as well as the complete COOH-terminus.

demonstrated that the AQP0 COOH-terminus is crucial for both lens development and transparency because of its interactions with calmodulin [20], the cytoskeletal proteins filensin and CP49 [21], and connexin 45.6 [22]. Cleavage of the intracellular COOH-terminus decreases water permeability [23] and enhances the adhesive properties of the extracellular surface of AQP0, indicating a conformational change in the molecule [24] such that AQP0-LHS may fail to form a functional channel. As a result, the internal homeostasis of the lens necessary to maintain transparency is disrupted, resulting in cataract formation.

In conclusion, analysis using an in vitro minigene system demonstrates that the *MIP* c.607-1G>A mutation leads to aberrant splicing by activation of a cryptic acceptor splice site in the 3' UTR of the *MIP* gene. To our knowledge, this is the first report of cryptic acceptor splice site in the 3' UTR of a human disease gene. The c.607-1G>A mutation creates a de novo AQP0-LHS, which is predicted to cause congenital cataracts by disrupting the internal homeostasis of the lens fiber cells.

TABLE 1. ACTIVATION OF ABERRANT 3' SPLICE SITES IN TERMINAL INTRONS/EXONS.

Gene	Phenotype	Mutation	Location of cryptic 3' ss	Location of stop codon
<i>CLN3</i>	Batten disease	IVS15-1G>T	E15+7	E15+120
<i>G6PC</i>	Glycogen storage disease type 1a	E5+86G>T	E5+91	E5+512
<i>HPRT1</i>	Hypoxanthine-guanine phosphoribosyltransferase deficiency	IVS9-1G>A	E10+17	E10+48
<i>HPRT2</i>	Hypoxanthine-guanine phosphoribosyltransferase deficiency	IVS9-2A>G	E10+17	E10+48
<i>HPRT3</i>	Hypoxanthine-guanine phosphoribosyltransferase deficiency	IVS9-2A>T	E10+17	E10+48
<i>CD40LG</i>	X-linked hyper-IgM syndrome	IVS4-2A>G	E5+8	E5+377
<i>UGT1A1</i>	Crigler-Najjar syndrome type 1	IVS4-1G>A	E5+7	E5+298
<i>APOE</i>	ApoE deficiency	IVS3-2A>G	IVS3-52	E4+718
<i>ERCC3</i>	Xeroderma pigmentosum	IVS14-6C>A	IVS14-4	E15+132
<i>HBB</i>	Beta-thalassaemia	IVS2-A>G	IVS2-271	E3+129
<i>LICAM</i>	X-linked hydrocephalus, MASA syndrome, spastic paraplegia	IVS26-12G>A	IVS26-10	E27+232
<i>MLYCD</i>	Malonyl-CoA decarboxylase deficiency	IVS4-14A>G	IVS4-13	E5+534
<i>MPO</i>	Myeloperoxidase deficiency	IVS11-2A>C	IVS11-109	E12+208
<i>SALL1</i>	Townes-Brocks syndrome	IVS2-19T>A	IVS2-17	E3+441
<i>SOD1</i>	Amyotrophic lateral sclerosis	IVS4-10T>G	IVS4-9	E5+108
<i>TPMT</i>	Thiopurine methyltransferase deficiency	IVS9-1G>A	IVS9+1, IVS9-330	E10+113

ACKNOWLEDGMENTS

We are grateful to Professor J. Fielding Hejtmancik at the Molecular Ophthalmic Genetics Section, Ophthalmic Genetics and Visual Function Branch, National Eye Institution, National Institutes of Health for his critical reading of this manuscript. This work was supported by Key Projects in the National Science & Technology Pillar Program in the Eleventh Five-year Plan Period (2006BAI02B04).

REFERENCES

- Amaya L, Taylor D, Russell-Eggitt I, Nischal KK, Lengyel D. The morphology and natural history of childhood cataracts. *Surv Ophthalmol* 2003; 48:125-44. [PMID: 12686301]
- Reddy MA, Francis PJ, Berry V, Bhattacharya SS, Moore AT. Molecular genetic basis of inherited cataract and associated phenotypes. *Surv Ophthalmol* 2004; 49:300-15. [PMID: 15110667]
- Hejtmancik JF. Congenital cataracts and their molecular genetics. *Semin Cell Dev Biol* 2008; 19:134-49. [PMID: 18035564]
- Graw J. Genetics of crystallins: cataract and beyond. *Exp Eye Res* 2009; 88:173-89. [PMID: 19007775]
- Berry V, Francis P, Kaushal S, Moore A, Bhattacharya S. Missense mutations in MIP underlie autosomal dominant 'polymorphic' and lamellar cataracts linked to 12q. *Nat Genet* 2000; 25:15-7. [PMID: 10802646]
- Gu F, Zhai H, Li D, Zhao L, Li C, Huang S, Ma X. A novel mutation in major intrinsic protein of the lens gene (MIP) underlies autosomal dominant cataract in a Chinese family. *Mol Vis* 2007; 13:1651-6. [PMID: 17893667]
- Lin H, Hejtmancik JF, Qi Y. A substitution of arginine to lysine at the COOH-terminus of MIP caused a different binocular phenotype in a congenital cataract family. *Mol Vis* 2007; 13:1822-7. [PMID: 17960133]
- Wang W, Jiang J, Zhu Y, Li J, Jin C, Shentu X, Yao K. A novel mutation in the major intrinsic protein (MIP) associated with autosomal dominant congenital cataracts in a Chinese family. *Mol Vis* 2010; 16:534-9. [PMID: 20361015]
- Geyer DD, Spence MA, Johannes M, Flodman P, Clancy KP, Berry R, Sparkes RS, Jonsen MD, Isenberg SJ, Bateman JB. Novel single-base deletional mutation in major intrinsic protein (MIP) in autosomal dominant cataract. *Am J Ophthalmol* 2006; 141:761-3. [PMID: 16564824]
- Jiang J, Jin C, Wang W, Tang X, Shentu X, Wu R, Wang Y, Xia K, Yao K. Identification of a novel splice-site mutation in MIP in a Chinese congenital cataract family. *Mol Vis* 2009; 15:38-44. [PMID: 19137077]
- Lareau LF, Green RE, Bhatnagar RS, Brenner SE. The evolving roles of alternative splicing. *Curr Opin Struct Biol* 2004; 14:273-82. [PMID: 15193306]
- Královicová J, Christensen MB, Vorechovsky I. Biased exon/intron distribution of cryptic and de novo 3' splice sites. *Nucleic Acids Res* 2005; 33:4882-98. [PMID: 16141195]
- Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 1997; 18:2714-23. [PMID: 9504803]
- Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 2003; 31:3381-5. [PMID: 12824332]
- Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 2006; 22:195-201. [PMID: 16301204]
- Vorechovský I. Aberrant 3' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 2006; 34:4630-41. [PMID: 16963498]
- Okada H, Kunishima S, Hamaguchi M, Takagi A, Yamamoto K, Takamatsu J, Matsushita T, Saito H, Kojima T, Yamazaki T. A novel splice site mutation in intron C of PROS1 leads to markedly reduced mutant mRNA level, absence of thrombin-

- sensitive region, and impaired secretion and cofactor activity of mutant protein S. *Thromb Res* 2010; 125:e246-50. [PMID: 20022358]
18. Cooper TA. Use of minigene systems to dissect alternative splicing elements. *Methods* 2005; 37:331-40. [PMID: 16314262]
 19. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009; 37:e67. [PMID: 19339519]
 20. Rose KM, Wang Z, Magrath GN, Hazard ES, Hildebrandt JD, Schey KL. Aquaporin 0-calmodulin interaction and the effect of aquaporin 0 phosphorylation. *Biochemistry* 2008; 47:339-47. [PMID: 18081321]
 21. Lindsey Rose KM, Gourdie RG, Prescott AR, Quinlan RA, Crouch RK, Schey KL. The C-terminus of lens aquaporin 0 interacts with the cytoskeletal proteins filensin and CP49. *Invest Ophthalmol Vis Sci* 2006; 47:1562-70. [PMID: 16565393]
 22. Yu XS, Yin X, Lafer EM, Jiang JX. Developmental regulation of the direct interaction between the intracellular loop of connexin 45.6 and the C-terminus of major intrinsic protein (aquaporin-0). *J Biol Chem* 2005; 280:22081-90. [PMID: 15802270]
 23. Ball LE, Little M, Nowak MW, Garland DL, Crouch RK, Schey KL. Water permeability of C-terminally truncated aquaporin 0 (AQP0 1-243) observed in the aging human lens. *Invest Ophthalmol Vis Sci* 2003; 44:4820-8. [PMID: 14578404]
 24. Gonen T, Sliz P, Kistler J, Cheng Y, Walz T. Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature* 2004; 429:193-7. [PMID: 15141214]