



Published in final edited form as:

*Hum Genet.* 2013 July ; 132(7): 761–770. doi:10.1007/s00439-013-1289-0.

## Whole Exome Sequencing in Dominant Cataract Identifies a New Causative Factor, *CRYBA2*, and a Variety of Novel Alleles in Known Genes

Linda M. Reis<sup>1</sup>, Rebecca C. Tyler<sup>1</sup>, Sanaa Muheisen<sup>1</sup>, Victor Raggio<sup>2</sup>, Leonardo Salviati<sup>3</sup>, Dennis P. Han<sup>4</sup>, Deborah Costakos<sup>4</sup>, Hagith Yonath<sup>5</sup>, Sarah Hall<sup>6</sup>, Patricia Power<sup>7</sup>, and Elena V. Semina<sup>1,8,\*</sup>

<sup>1</sup>Department of Pediatrics and Children's Research Institute, Medical College of Wisconsin, Milwaukee, WI, 53226 USA

<sup>2</sup>Genetics Department, School of Medicine, Montevideo, PC: 11600, Uruguay

<sup>3</sup>Dept. of Woman and Child Health, University of Padova, 35128 Padova, Italy

<sup>4</sup>Department of Ophthalmology, Medical College of Wisconsin, Milwaukee, WI 53226 USA

<sup>5</sup>Sheba Medical Center, Tel Hashomer 52621 and Sackler School of Medicine, Tel Aviv University, Israel

<sup>6</sup>Kadlec Regional Medical Center, Richland WA 99352 USA

<sup>7</sup>Department of Medical Genetics, University of British Columbia, Vancouver, BC V6H 3N1 Canada

<sup>8</sup>Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI, 53226 USA

### Abstract

Pediatric cataracts are observed in 1–15 per 10,000 births with 10–25% of cases attributed to genetic causes; autosomal dominant inheritance is the most commonly observed pattern. Since the specific cataract phenotype is not sufficient to predict which gene is mutated, whole exome sequencing (WES) was utilized to concurrently screen all known cataract genes and to examine novel candidate factors for a disease-causing mutation in probands from 23 pedigrees affected with familial dominant cataract. Review of WES data for 36 known cataract genes identified causative mutations in nine pedigrees (39%) in *CRYAA*, *CRYBB1*, *CRYBB3*, *CRYGC* (2), *CRYGD*, *GJA8* (2), and *MIP* and an additional likely causative mutation in *EYA1*; the *CRYBB3* mutation represents the first dominant allele in this gene and demonstrates incomplete penetrance. Examination of crystallin genes not yet linked to human disease identified a novel cataract gene, *CRYBA2*, a member of the  $\beta\gamma$ -crystallin superfamily. The p.(Val50Met) mutation in *CRYBA2* cosegregated with disease phenotype in a four-generation pedigree with autosomal dominant congenital cataracts with incomplete penetrance. Expression studies detected *cryba2* transcripts during early lens development in zebrafish, supporting its role in congenital disease. Our data highlight the extreme genetic heterogeneity of dominant cataract as the eleven causative/likely causative mutations affected nine different genes and the majority of mutant alleles were novel.

\*Correspondence: Phone: 414-955-4996; Fax: 414-955-6329; esemina@mcw.edu.

#### Ethical standards

All experiments described comply with the current laws of the United States of America.

#### Conflict of interest

The authors declare that they have no conflict of interest.

Furthermore, these data suggest that less than half of dominant cataract can be explained by mutations in currently known genes.

## Keywords

dominant cataract; whole exome sequencing; *CRYBA2*; *CRYBB3*; zebrafish

---

## Introduction

Pediatric cataracts are observed in 1–15 per 10,000 births, with 10–25% of cases attributed to genetic causes (Trumler 2011). Autosomal dominant inheritance is most common for hereditary cases, although autosomal recessive and X-linked patterns are observed (Hejtmancik 2008). Congenital or infantile cataract represents the most severe end of the spectrum with onset within the first year of life and the possibility of permanent blindness due to interference with normal visual development. Genetic studies have identified mutations in numerous genes associated with cataracts, including crystallins, which comprise about half of the known genetic forms of cataract, structural proteins, membrane proteins, and transcription factors (Hejtmancik 2008; Huang and He 2010; Shiels et al. 2010). The exact frequencies of mutations in these identified genes are not known and many cases are still awaiting molecular diagnosis.

Whole exome sequencing has been successfully applied for molecular characterization of a cohort affected with a heterogeneous condition (Aldahmesh et al. 2012; Choi et al. 2012) as well as identification of novel genes (Bamshad et al. 2011). Since the specific cataract phenotype is typically not sufficient to predict which gene is mutated in a family (Hejtmancik 2008), whole exome sequencing may represent an efficient method of screening the known cataract genes in order to identify a disease-causing mutation as well as looking for novel factors involved in this condition. We utilized whole exome sequencing in a population affected with dominant cataract to first provide insight regarding the spectrum/frequency of mutations in known genes and the proportion of hereditary cataract which remains unexplained and second, to perform a comprehensive examination of crystallin genes not previously associated with human disease for mutations in cataract phenotypes.

## Methods

### Ethics statement

The human study was approved by the Institutional Review Board of the Children's Hospital of Wisconsin with written informed consent obtained from each participant and/or their legal representative, as appropriate. The studies involving zebrafish embryos were carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Institutional Animal Care and Use Committee at the Medical College of Wisconsin (protocol number AUA00000352).

### Subjects

Probands from 23 pedigrees affected with dominant cataract were selected for whole exome sequencing: 18 had cataract diagnosed within the first year of life in at least one member of the family (congenital cataract), 4 were diagnosed between 1 and 5 years of age (juvenile cataract), and one family demonstrated presenile adult-onset (age 40) cataract and retinal detachment (Online Resources 1 and 2). Two pedigrees with cataract microcornea and previously identified mutations in *PITX3/FOXE3* were excluded from this study (data not

shown). The 23 pedigrees included 19 Caucasian (European and European American), 3 Hispanic, and 1 family with unreported race/ethnicity. Additional family members were available for testing in 21 cases, ranging from 1–12 additional individuals (Online Resource 1).

### Whole Exome Sequencing and Data Analysis

Whole exome sequencing was undertaken through Perkin Elmer, Inc (Branford, CT). Patients 1–14 utilized Agilent Sure Select v4 for exome capture while Patients 15–23 utilized v4+UTR. Data was evaluated through the Geospiza GeneSifter Analysis program hosted through Perkin Elmer Bioinformatics using the GATK V2.10 pipeline. The Variant Report was reviewed for the presence of mutations in 36 known cataract genes and 8 additional crystallin genes not yet linked to cataract in humans (Online Resources 2 and 3) while the Gene List was used to verify coverage of the genes of interest. Variants of interest were investigated for presence/absence in NCBI SNP Database (dbSNP) (<http://www.ncbi.nlm.nih.gov/projects/SNP/>), frequency in the Exome Variant Server (EVS) (<http://evs.gs.washington.edu/EVS/>), and predicted effect on the protein (Ensembl Variant Effect Predictor: <http://www.ensembl.org/info/docs/variation/vep/index.html>).

### Variant Confirmation

Primers flanking the variants of interest were used to amplify genomic DNA from probands (to confirm variant) and all available family members (to test for cosegregation). PCR products were sequenced in both directions using ABI 3730XL sequencer and protocols (Applied Biosystems/Life Technologies, Carlsbad, CA, USA). Sequences were reviewed manually and using Mutation Surveyor (SoftGenetics, State College, PA).

### Mutation Analysis of *MAF*, *FOXE3*, and *PITX3*

*MAF*, *FOXE3* and *PITX3* were analyzed by direct DNA sequencing of PCR products in all probands following previously described protocols (Bremond-Gignac et al. 2010; Hansen et al. 2007a).

### In situ hybridization in zebrafish embryos

Zebrafish (*Danio rerio*) were maintained on a 14-hour light/ 10-hour dark cycle. The embryos were obtained by natural spawning and maintained at 28.5°C. Expression analysis of two zebrafish orthologs of human *CRYBA2*, *cryba2a* (NM\_001002049.1) and *cryba2b* (NM\_001002584.1), was performed by in situ hybridization with 768-nt (*cryba2a*) and 752-nt (*cryba2b*) transcript-specific antisense riboprobes and previously described protocols (Sorokina et al. 2011). To generate templates for probe making, the corresponding *cryba2a* and *cryba2b* transcripts were amplified using the specific primers, *cryba2a* forward, GCCAAATCTCTCCCACGACA, and reverse, GAATGGCGACAAGCACACTC, as well as *cryba2b* forward, GCATTCGCCACTGAATGAGG, and reverse, TGCCTATAGTATTGATACGG, and cloned into pCRII-TOPO vector containing T7 and Sp6 promoters for in vitro RNA synthesis (Invitrogen, Carlsbad, CA).

## Results/Discussion

### Analysis of thirty-six known genes associated with pediatric cataract

The average mean read depth for the whole exome was 68.5 and on average, 86% of the targeted exome region achieved coverage >10X (Online Resource 2). Analysis of whole exome data for the 36 known cataract genes (Online Resource 3) showed generally good coverage. Three genes (*MAF*, *FOXE3* and *PITX3*) showed poor coverage via the Agilent Sure Select v4 capture kit with average coding region coverage of 7×; coverage of these

three genes was improved via Agilent Sure Select v4+UTR capture kit with average coding region coverage of 53×. *MAF*, *FOXE3* and *PITX3* sequencing was completed by Sanger sequencing in all probands. The remaining 33 cataract genes showed good coverage via both v4 (average coding region coverage of 71×) and v4+UTR (average coding region coverage of 64×) capture kits.

Heterozygous causative mutations in known cataract genes were identified in nine families in *CRYAA*, *CRYBB1*, *CRYBB3*, *CRYGC* (two mutations), *CRYGD*, *GJA8* (two mutations), and *MIP*, and an additional likely causative mutation was seen in *EYAI* (Table 1). Three of the mutations represent new occurrences of previously reported mutations (*CRYAA*, *CRYGC* and *CRYGD*); the remaining seven are novel. None of these mutations were present in dbSNP or the EVS. All nine causative mutations in known genes were identified in families affected with congenital cataract; each of these mutations showed complete cosegregation with the disease phenotype, with incomplete penetrance noted in one family (Table 1). The majority of the identified mutations are consistent with previous reports; two changes are of particular interest, the mutation in *CRYBB3* that identifies a novel inheritance pattern for this gene and a likely causative change in *EYAI* associated with an isolated cataract phenotype.

A novel heterozygous missense mutation in *CRYBB3*, c.581T>A, p.(Val194Glu), was identified in Patient 7 from an extended pedigree affected with bilateral congenital cataract demonstrating dominant inheritance with reduced penetrance (Figure 1). The specific types of cataract varied (posterior polar (IV-4, IV-6), nuclear (IV-5, and V-5), or anterior polar plus cortical (V-4)) and two individuals also had glaucoma (V-4 and V-5); cataracts were diagnosed at birth in four cases and at one year of age in one (IV-5). The two obligate carriers, III-2 and III-3, were last examined by an ophthalmologist at ages 46 and 70, correspondingly, and had no evidence of cataract. The missense mutation is seen in all affected individuals tested (IV-4, IV-5, IV-6, V-4, V-5) as well as an unaffected sibling (IV-3) (last ophthalmological evaluation occurred at age 35 and was reported normal); unfortunately, no other unaffected family members were available for genetic testing (Figure 1). The c.581T>A/ p.(Val194Glu) mutation was concluded to be causative because it cosegregates with affected status in multiple individuals from two branches of the family, is not seen in over 10,000 chromosomes screened by EVS, and results in a nonconservative substitution of a highly conserved amino acid valine within the Greek key motif IV of *CRYBB3*, which is predicted to be probably damaging/deleterious by PolyPhen/SIFT (accessed through Ensembl Variant Effect Predictor).

While the previous mutation reported in *CRYBB3* was recessive (Riazuddin et al. 2005), all other crystallins are associated with autosomal dominant inheritance and three, *CRYAB*, *CRYAA* and *CRYBB1* are also associated with both recessive and dominant inheritance, along with *GJA8*, *SIL1*, *HSF4*, and *EPHA2* (Huang and He 2010; Kaul et al. 2010). For several genes, there seems to be a clear correlation between the type/location of a mutation and its pattern of inheritance, with nonsense mediated decay (NMD) likely playing a role in some cases (see below). As for *CRYBB3*, the only previously reported mutation in this gene is the recessive allele c.493G>C, p.(Gly165Arg), located in the loop connecting  $\beta$ -strands 1 and 2 of the Greek-key motif IV (Riazuddin et al. 2005), while the dominant allele reported here, c.581T>A, p.(Val194Glu), occurs within the  $\beta$ -strand 4 of the same motif (Slingsby and Clout 1999). Both changes are nonconservative, with the p.(Val194Glu) mutation causing the hydrophobic amino acid valine to be replaced with hydrophilic glutamic acid, therefore likely resulting in destabilization of the domain structure of the Greek key motif IV and *CRYBB3* protein folding. Since this is only the second mutation identified in *CRYBB3*, additional studies are needed to determine genotype-phenotype correlations and differences in the mechanisms of these mutations. It is also possible that the dominance/

recessivity of *CRYBB3* alleles can be modified by secondary factors, consistent with the incomplete penetrance of the dominant c.581T>A, p.(Val194Glu) allele reported in this manuscript. Incomplete penetrance has been associated with mutations in several other autosomal dominant cataract genes: *GJA8* (He et al. 2011), *FTL* (Gonzalez-Huerta et al. 2008), *GJA3* (Burdon et al. 2004), *CRYGC* (Ren et al. 2000), and *CRYAB* (Sacconi et al. 2012).

The nucleotide change in *EYAI* observed in Patient 15 results in a nonconservative substitution of the glutamic acid for lysine in the N-terminal region of this regulatory protein. *EYAI* mutations are strongly associated with Brachio-Oto-Renal (BOR) syndrome, with no particular genotype-phenotype correlation identified between the type/location of mutation and phenotype severity (Orten et al. 2008). *EYAI* mutations in isolated cataract/anterior segment dysgenesis phenotypes are rare and so far documented by only one published report that described two missense changes (Azuma et al. 2000). The c.121G>A, p.(Glu41Lys) substitution observed in Patient 15 with isolated congenital cataract is predicted to be possibly damaging by PolyPhen and is not observed in over 12,000 chromosomes screened by EVS, which suggests that this allele may be causative for the cataract phenotype observed in this patient. Unfortunately, samples from other family members were not available to verify segregation of the c.121G>A, p.(Glu41Lys) allele with the affected phenotype and further confirm (or refute) its pathogenicity.

An additional 14 heterozygous novel or rare (<1% population frequency) variants were identified in 11 cataract genes that did not segregate with the disease phenotype in affected pedigrees and therefore were determined to be non-causative (Online Resource 4); two findings in autosomal dominant genes were of special interest. A rare missense variant in *GJA8*, c.741T>G, p.(Ile247Met), was seen in Patient 2 (with causative *CRYBB1* mutation) but did not cosegregate with the disease phenotype. This allele was initially reported as a disease causing mutation in a small family affected with congenital cataracts (Polyakov et al. 2001), although a subsequent report suggested it was a rare polymorphism (Graw et al. 2009) and it is present in 0.7% of the European American population in the EVS (31/8600 chromosomes); our data provide further evidence that the c.741T>G, p.(Ile247Met) variant in *GJA8* is a polymorphism.

The second variant of interest was a novel nonsense allele in *CRYGD*, c.51T>G, p.(Tyr17\*), seen in Patient 23 with adult-onset cataract and not reported in almost 13,000 chromosomes screened in the EVS; the variant did not cosegregate with the disease phenotype, being present in the patient, an affected sibling, and an unaffected sibling but not seen in an affected sibling and their mother. At the same time, four other nonsense mutations in *CRYGD* (p.(Tyr56\*), p.(Tyr134\*), p.(Arg140\*) and (p.Trp156\*)) have been linked to pediatric cataract phenotypes (Huang and He 2010), including an occurrence in Patient 6 from this study. The difference in effects generated by these nonsense mutations is likely to be due to whether or not they are subjected to nonsense mediated decay (NMD) (Holbrook et al. 2004; Khajavi et al. 2006). All previously reported *CRYGD* dominant nonsense mutations occur later in the gene; three out of four, p.(Tyr134\*), p.(Arg140\*) and (p.Trp156\*), are located the final exon and thus are predicted to escape NMD resulting in truncated protein products with altered structural conformation, which may generate a dominant-negative effect (Hansen et al. 2007b). In contrast, the c.51T>G, p.(Tyr17\*), *CRYGD* variant reported here occurs in the beginning of the second exon and is predicted to be subject to NMD and not expected to yield a protein product; this mutation might result in autosomal recessive disease if it occurred as a homozygous (or compound heterozygous) mutation. Based on this data, the previously reported dominant *CRYGD* allele, nonsense mutation p.(Tyr56\*), may need to be re-examined. While the mutation was reported in a dominant cataract family (Santana et al. 2009), the mutation is located in the second exon

more than 55 nucleotides from the final intron and thus would be expected to be subject to NMD, similar to the p.(Tyr17\*) variant reported here. NMD alters the patterns of inheritance for premature truncation alleles in many genes with 5' nonsense mutations (predicted to be subject to NMD) resulting in recessive disease and 3' nonsense mutations (predicted to escape NMD) leading to dominant disease (Khajavi et al. 2006). In terms of cataract genes, examples include *CRYBB1* mutations with dominant changes comprising missense mutations in the Greek key motifs II or IV, a C-terminal extension mutation, and C-terminal truncations that retain ~90% of normal protein sequence but lack part of the Greek key motif IV (all located in the final *CRYBB1* exon and predicted to escape NMD); all of these changes are likely to result in production of mutant proteins. In contrast, recessive *CRYBB1* changes include an early frameshift mutation that is likely to be subject to NMD, and an initiation codon substitution, both most likely resulting in null alleles (Cohen et al. 2007; Meyer et al. 2009; Wang et al. 2011). Similarly, a nonsense mutation early in the *CRYAA* gene, p.(Trp9\*), is associated with autosomal recessive cataract while most other mutations in this gene result in dominant disease (Graw 2009; Pras et al. 2000). Our analysis of genetic variants highlights the importance of screening all known genes associated with a phenotype and completing cosegregation analysis in the pursuit of a causative mutation in a given pedigree.

### Analysis of crystallin genes not yet linked to human disease

Since crystallin mutations comprised the majority of causative alleles (6/9), we proceeded to analyze the sequence of eight crystallin genes not previously linked to cataract in humans: *CRYGB*, *CRYBA2*, *CRYL1*, *CRYGN*, *CRYZ*, *CRYM*, *CRYZL1*, and *CRYBG3* (Online Resource 2).

This analysis identified a novel heterozygous missense mutation, c.148G>A, p.(Val50Met), in the *CRYBA2* gene in Patient 4 (Table 1, Figure 2). Patient 4 is affected with bilateral multifocal congenital cataract and eccentric pupil. He has a strong family history of autosomal dominant cataracts with incomplete penetrance (Figure 2); glaucoma and/or myopia were also noted in some affected individuals. The cataracts were congenital in all cases, with the exception of individual III-10, whose cataracts were diagnosed at age 9 and individual III-18, whose cataracts were diagnosed at age 34, after the birth of an affected child; it is possible that these individuals may have had mild congenital disease which went undiagnosed due to incomplete evaluation earlier in life. Three unaffected family members who are obligate carriers based on the pedigree were also available for testing. All seven affected individuals tested (III-3, III-8, III-10, III-14, III-18) as well as the obligate carriers (II-1, III-2, and III-6) carry the missense mutation while the unaffected spouses (II-2 and III-7) and an unaffected child (IV-5) do not, providing strong evidence that this mutation is causative of the phenotype in this family. In addition to this, the mutation affects a highly conserved valine residue within the  $\beta$ -strand 4 of the Greek key motif I that is preserved in all  $\beta\gamma$ -crystallins (Figure 2) and thus is predicted to be probably damaging/deleterious by PolyPhen/SIFT (accessed through Ensembl Variant Effect Predictor) and the mutant allele was not seen in almost 13,000 chromosomes reported by the EVS.

No other candidates for disease causing mutations were identified in this family (Online Resource 2). Three rare/novel variants in autosomal dominant genes *BFSP2* and *CRYGS* were seen on chromosome 3 in the proband, Patient 4 (Online Resource 4); these three changes did not segregate with the disease phenotype but appeared to travel as a haplotype within the family with the exception of individual IV-3, who had only the second two changes, and individual IV-4, who had only the first (data not shown).

Previous studies demonstrated *CRYBA2/Cryba2* expression in adult human lens epithelial and fiber cells (Hawse et al. 2005) as well as in mouse lens epithelial and fiber cells at

postnatal days 1, 7, and 21 (Puk et al. 2011). To further evaluate the role of *CRYBA2* in lens development, we studied expression of its zebrafish orthologs during embryonic development by in situ hybridization with *cryba2a*- and *cryba2b*- specific antisense riboprobes and detected a strong and specific presence of both *cryba2* transcripts at early stages of lens development (Figure 3): at 22-hpf, *cryba2* orthologs are expressed in the posterior region of the lens mass that delaminates from the head ectoderm (corresponding to lens vesicle formation/separation in 4-week-old human embryos (Smelser 1965) (Figure 3c); at 48-hpf, robust expression is observed in the developing lens fiber cells (Figure 3f). The strong expression of *cryba2* transcripts at early stages of lens development provides further support for its role in congenital disease.

The discovery of a causative mutation in *CRYBA2*, a member of  $\beta\gamma$ -crystallin superfamily, provides further evidence of the importance of these proteins to ocular function. Crystallins represent the major structural components of the vertebrate lens and function to increase lens refractory power while preserving its transparency. Mutations in crystallins are widely associated with cataracts in various species (Graw 2009). The  $\beta\gamma$ -crystallin superfamily is comprised of thirteen members in humans and ten of them have now been shown to cause cataracts in humans when mutated. In mouse, a missense mutation in *Cryba2*, c.139T>C, p.(Ser47Pro), was associated with small lenses and age-related cataracts at 25 weeks in both heterozygous and homozygous animals; the phenotype was more severe in homozygous mice, but no evidence of congenital cataract was seen in mice with either genotype with histologic analysis performed at postnatal days 1 and 21 (Puk et al. 2011). Interestingly, the disease-causing mutations in both mouse, p.(Ser47Pro), and human, p.(Val50Met), affect invariant amino acids within the Greek key motif I and are separated by only two residues. While it is possible that human eyes are more sensitive to *CRYBA2* mutations, other factors potentially contributing to the differences in the age of onset, such as the specific genetic background or differences in lifespan/development between species, cannot be ruled out. Identification of additional human and mouse *CRYBA2/cryba2* alleles will allow for better understanding of the phenotypic spectrums associated with mutations in this gene.

No disease-causing mutations were identified in the other seven novel crystallin genes screened (*CRYGB*, *CRYL1*, *CRYGN*, *CRYZ*, *CRYM*, *CRYZL1*, and *CRYBG3*). One rare variant, c.97C>T, p.(Arg33Trp), was seen in *CRYGN* in Patient 1 and is predicted to be benign/ deleterious by PolyPhen/SIFT but did not segregate with the disease phenotype; this variant was also seen in 2/8600 EA chromosomes in EVS (0.05% population frequency).

In summary, analysis of twenty-three dominant cataract pedigrees by whole exome sequencing identified causative mutations in known cataract genes (*CRYAA*, *CRYBB1*, *CRYBB3*, *CRYGC* (2), *CRYGD*, *GJA8* (2), and *MIP*) in nine families (39%), a novel causative gene mutation in a member of  $\beta\gamma$ -crystallin superfamily *CRYBA2* in one pedigree, and an additional likely causative mutation in *EYAI* in one case, thus explaining up to 48% (11/23) of dominant cataract in our study. All mutations were identified in families affected with congenital cataract (11/18); in contrast, no mutations were found in families affected with familial juvenile or adult-onset cataracts (0/5). Our data highlight the extreme genetic heterogeneity of dominant congenital cataract as the eleven causative/likely causative mutations affected nine different genes and the majority of mutant alleles were novel. Furthermore, these data suggest that more than half of hereditary dominant cataract remains to be explained. While it is possible that some of these families may have mutations in known genes in small regions not well-covered by WES or copy number variations affecting these genes, it is likely that the majority are due to novel factors. Ongoing analysis of WES data generated in this study will likely identify novel genetic factors responsible for the cataract seen in the remaining families.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors gratefully acknowledge the patients and their families for their participation in research studies. This work was supported by the National Institutes of Health awards R01EY015518, R21DC010912 and funds provided by the Children's Hospital of Wisconsin (EVS), 1UL1RR031973 from the Clinical and Translational Science Award (CTSA) program, and was supported in part by Research to Prevent Blindness, Inc., New York, NY (DPH and DC); the authors are also thankful to Elena A. Sorokina for cloning of zebrafish *cryba2* gene sequences and thoughtful discussions.

## References

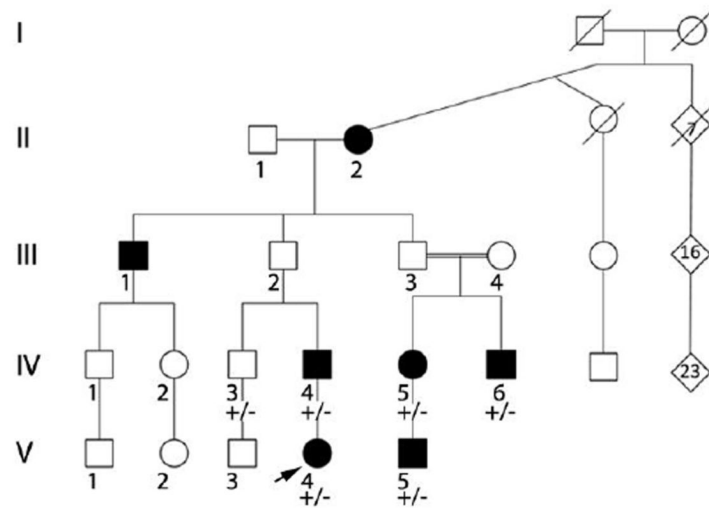
- Aldahmesh MA, Khan AO, Mohamed JY, Hijazi H, Al-Owain M, Alswaid A, Alkuraya FS. Genomic analysis of pediatric cataract in Saudi Arabia reveals novel candidate disease genes. *Genet Med*. 2012
- Azuma N, Hirakiyama A, Inoue T, Asaka A, Yamada M. Mutations of a human homologue of the *Drosophila* eyes absent gene (EYA1) detected in patients with congenital cataracts and ocular anterior segment anomalies. *Hum Mol Genet*. 2000; 9:363–366. [PubMed: 10655545]
- Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011; 12:745–755. [PubMed: 21946919]
- Bremond-Gignac D, Bitoun P, Reis LM, Copin H, Murray JC, Semina EV. Identification of dominant FOXE3 and PAX6 mutations in patients with congenital cataract and aniridia. *Mol Vis*. 2010; 16:1705–1711. [PubMed: 20806047]
- Burdon KP, Wirth MG, Mackey DA, Russell-Eggitt IM, Craig JE, Elder JE, Dickinson JL, Sale MM. A novel mutation in the Connexin 46 gene causes autosomal dominant congenital cataract with incomplete penetrance. *J Med Genet*. 2004; 41:e106. [PubMed: 15286166]
- Choi BO, Koo SK, Park MH, Rhee H, Yang SJ, Choi KG, Jung SC, Kim HS, Hyun YS, Nakhro K, Lee HJ, Woo HM, Chung KW. Exome sequencing is an efficient tool for genetic screening of Charcot-Marie-Tooth Disease. *Hum Mutat*. 2012; 33:1610–1615. [PubMed: 22730194]
- Cohen D, Bar-Yosef U, Levy J, Gradstein L, Belfair N, Ofir R, Joshua S, Lifshitz T, Carmi R, Birk OS. Homozygous CRYBB1 deletion mutation underlies autosomal recessive congenital cataract. *Invest Ophthalmol Vis Sci*. 2007; 48:2208–2213. [PubMed: 17460281]
- Gonzalez-Huerta L, Ramirez-Sanchez V, Rivera-Vega M, Messina-Baas O, Cuevas-Covarrubias S. A family with hereditary hyperferritinaemia cataract syndrome: evidence of incomplete penetrance and clinical heterogeneity. *Br J Haematol*. 2008; 143:596–598. [PubMed: 18710380]
- Graw J. Genetics of crystallins: cataract and beyond. *Exp Eye Res*. 2009; 88:173–189. [PubMed: 19007775]
- Graw J, Schmidt W, Minogue PJ, Rodriguez J, Tong JJ, Klopp N, Illig T, Ebihara L, Berthoud VM, Beyer EC. The GJA8 allele encoding CX50I247M is a rare polymorphism, not a cataract-causing mutation. *Mol Vis*. 2009; 15:1881–1885. [PubMed: 19756179]
- Hansen L, Eiberg H, Rosenberg T. Novel MAF mutation in a family with congenital cataract-microcornea syndrome. *Mol Vis*. 2007a; 13:2019–2022. [PubMed: 17982426]
- Hansen L, Yao W, Eiberg H, Kjaer KW, Baggesen K, Hejtmancik JF, Rosenberg T. Genetic heterogeneity in microcornea-cataract: five novel mutations in CRYAA, CRYGD, and GJA8. *Invest Ophthalmol Vis Sci*. 2007b; 48:3937–3944. [PubMed: 17724170]
- Hawse JR, DeAmicis-Tress C, Cowell TL, Kantorow M. Identification of global gene expression differences between human lens epithelial and cortical fiber cells reveals specific genes and their associated pathways important for specialized lens cell functions. *Mol Vis*. 2005; 11:274–283. [PubMed: 15851978]



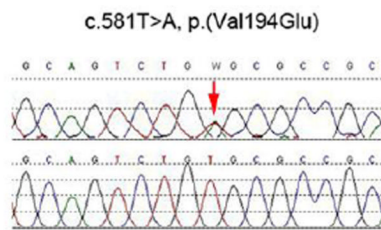
- He W, Li X, Chen J, Xu L, Zhang F, Dai Q, Cui H, Wang DM, Yu J, Hu S, Lu S. Genetic linkage analyses and Cx50 mutation detection in a large multiplex Chinese family with hereditary nuclear cataract. *Ophthalmic Genet.* 2011; 32:48–53. [PubMed: 21174522]
- Hejtmancik JF. Congenital cataracts and their molecular genetics. *Semin Cell Dev Biol.* 2008; 19:134–149. [PubMed: 18035564]
- Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. Nonsense-mediated decay approaches the clinic. *Nat Genet.* 2004; 36:801–808. [PubMed: 15284851]
- Huang B, He W. Molecular characteristics of inherited congenital cataracts. *Eur J Med Genet.* 2010; 53:347–357. [PubMed: 20624502]
- Kaul H, Riazuddin SA, Shahid M, Kousar S, Butt NH, Zafar AU, Khan SN, Husnain T, Akram J, Hejtmancik JF, Riazuddin S. Autosomal recessive congenital cataract linked to EPHA2 in a consanguineous Pakistani family. *Mol Vis.* 2010; 16:511–517. [PubMed: 20361013]
- Khajavi M, Inoue K, Lupski JR. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur J Hum Genet.* 2006; 14:1074–1081. [PubMed: 16757948]
- Meyer E, Rahman F, Owens J, Pasha S, Morgan NV, Trembath RC, Stone EM, Moore AT, Maher ER. Initiation codon mutation in betaB1-crystallin (CRYBB1) associated with autosomal recessive nuclear pulverulent cataract. *Mol Vis.* 2009; 15:1014–1019. [PubMed: 19461930]
- Orten DJ, Fischer SM, Sorensen JL, Radhakrishna U, Cremers CW, Marres HA, Van Camp G, Welch KO, Smith RJ, Kimberling WJ. Branchio-oto-renal syndrome (BOR): novel mutations in the EYA1 gene, and a review of the mutational genetics of BOR. *Hum Mutat.* 2008; 29:537–544. [PubMed: 18220287]
- Polyakov AV, Shagina IA, Khlebnikova OV, Evgrafov OV. Mutation in the connexin 50 gene (GJA8) in a Russian family with zonular pulverulent cataract. *Clin Genet.* 2001; 60:476–478. [PubMed: 11846744]
- Pras E, Frydman M, Levy-Nissenbaum E, Bakhan T, Raz J, Assia EI, Goldman B, Pras E. A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. *Invest Ophthalmol Vis Sci.* 2000; 41:3511–3515. [PubMed: 11006246]
- Puk O, Ahmad N, Wagner S, Hrabe de Angelis M, Graw J. First mutation in the betaA2-crystallin encoding gene is associated with small lenses and age-related cataracts. *Invest Ophthalmol Vis Sci.* 2011; 52:2571–2576. [PubMed: 21212184]
- Ren Z, Li A, Shastry BS, Padma T, Ayyagari R, Scott MH, Parks MM, Kaiser-Kupfer MI, Hejtmancik JF. A 5-base insertion in the gammaC-crystallin gene is associated with autosomal dominant variable zonular pulverulent cataract. *Hum Genet.* 2000; 106:531–537. [PubMed: 10914683]
- Riazuddin SA, Yasmeen A, Yao W, Sergeev YV, Zhang Q, Zulfiqar F, Riaz A, Riazuddin S, Hejtmancik JF. Mutations in betaB3-crystallin associated with autosomal recessive cataract in two Pakistani families. *Invest Ophthalmol Vis Sci.* 2005; 46:2100–2106. [PubMed: 15914629]
- Sacconi S, Feasson L, Antoine JC, Pecheux C, Bernard R, Cobo AM, Casarin A, Salviati L, Desnuelle C, Urtizberea A. A novel CRYAB mutation resulting in multisystemic disease. *Neuromuscul Disord.* 2012; 22:66–72. [PubMed: 21920752]
- Santana A, Waiswol M, Arcieri ES, Cabral de Vasconcellos JP, Barbosa de Melo M. Mutation analysis of CRYAA, CRYGC, and CRYGD associated with autosomal dominant congenital cataract in Brazilian families. *Mol Vis.* 2009; 15:793–800. [PubMed: 19390652]
- Shiels A, Bennett TM, Hejtmancik JF. Cat-Map: putting cataract on the map. *Mol Vis.* 2010; 16:2007–2015. [PubMed: 21042563]
- Slingsby C, Clout NJ. Structure of the crystallins. *Eye (Lond).* 1999; 13((Pt 3b)):395–402. [PubMed: 10627816]
- Smelser GK. Embryology and Morphology of the Lens. *Invest Ophthalmol.* 1965; 4:398–410. [PubMed: 14340157]
- Sorokina EA, Muheisen S, Mlodik N, Semina EV. MIP/Aquaporin 0 represents a direct transcriptional target of PITX3 in the developing lens. *PLoS One.* 2011; 6:e21122. [PubMed: 21698120]
- Trumler AA. Evaluation of pediatric cataracts and systemic disorders. *Curr Opin Ophthalmol.* 2011; 22:365–379. [PubMed: 21832913]
- Wang KJ, Wang S, Cao NQ, Yan YB, Zhu SQ. A novel mutation in CRYBB1 associated with congenital cataract-microcornea syndrome: the p.Ser129Arg mutation destabilizes the betaB1/

betaA3-crystallin heteromer but not the betaB1-crystallin homomer. *Hum Mutat.* 2011; 32:E2050–E2060. [PubMed: 21972112]

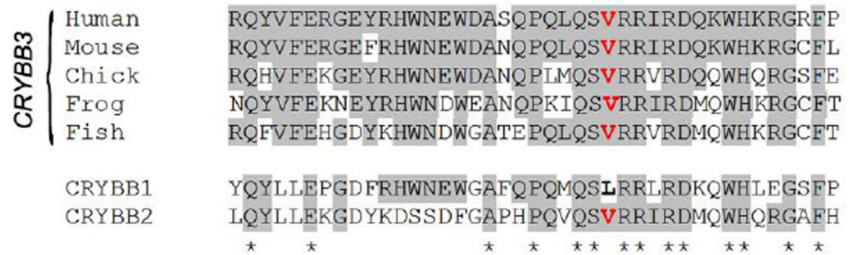
**a** Pedigree and genotyping data for *CRYBB3* mutation



**b** DNA sequence chromatogram

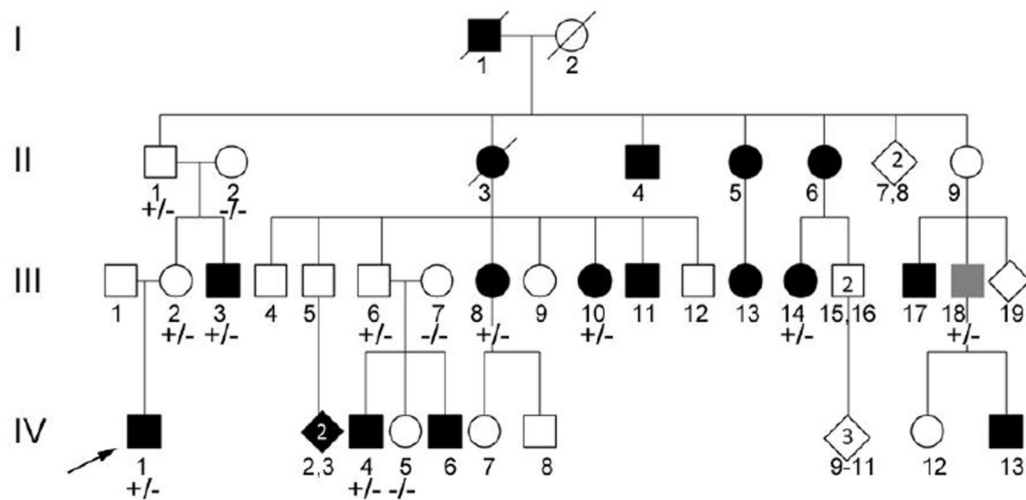


**c** Alignment of crystallin protein sequences

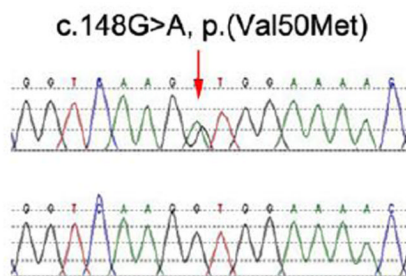


**Figure 1.** Pedigree, DNA sequence and amino acid conservation associated with *CRYBB3* allele. **a.** Pedigree of family 7 with affected status and genotyping results indicated (+/-: heterozygous *CRYBB3* mutation present). The proband is indicated with an arrow; black symbol indicates congenital or juvenile cataract. **b.** DNA sequence of *CRYBB3* showing heterozygous c.581T>A mutation (control sequence on bottom). **c.** Alignment of protein regions showing conservation of Val194 residue in the *CRYBB3* proteins in various species as well as in *CRYBB2*, but not *CRYBB1*.

### a Pedigree and genotyping data for *CRYBA2* mutation



### b DNA sequence chromatogram

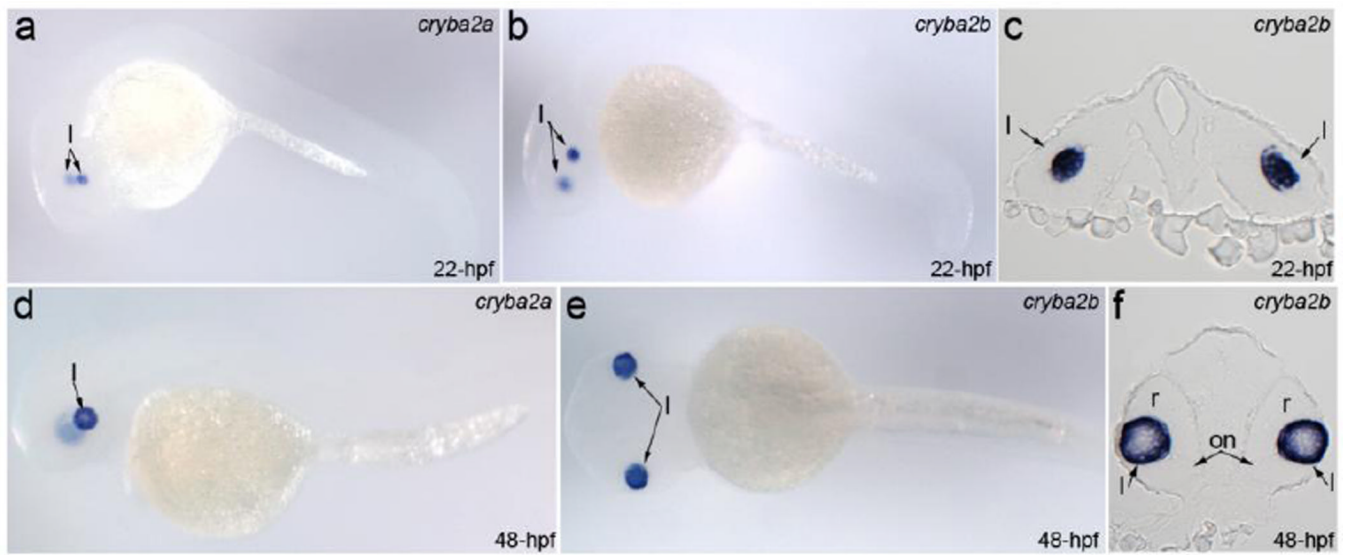


### c Alignment of crystallin protein sequences

CRYBA2	Human	NVCERGG LPRVRSVKV <b>V</b> ENG V WVA FEY P DF QGQQ F ILEKGDYPR
	Mouse	NVCERGA LRRVRSVKV <b>V</b> ENGA WVA FEY P DF QGQQ F ILEKGDYPC
	Chick	SIMERGF RK- I R S I K <b>V</b> E S G P W V G F E Y P E Y QGQQ F ILEKGDYPR
	Frog	NIMERGF QR- I R S I K <b>V</b> E N G P W V G F E Y P E F QGQQ F ILEKGDYPR
	Fish	NILDRDF QK- I R S I K <b>V</b> D N G P W V G Y E Y P E Y QGQQ F ILEKGDYPC
	CRYBA4	SVLELG FET- V R S L K <b>V</b> L S G A W V G F E H A G F QGQQ Y I L R G E Y P R
	CRYBA1	NV S E R- S F D N V R S L K <b>V</b> E S G A W I G Y E H T S F C G Q Q F I L E R G E Y P R
	CRYBB1	N L A D R G- F D R V R S I I <b>V</b> S A G P W V A F E Q S N F R G E M F I L E K G E Y P R
	CRYBB2	N L K E T G- V E K A G S V L <b>V</b> Q A G P W V G Y E Q A N C K G E Q F V F E K G E Y P R
	CRYBB3	P S L T D S L L E K V G S I Q <b>V</b> E S G P W L A F E S R A F R G E Q F V L E K G D Y P R
	CRYGC	C P N L Q P Y F S R C N S I R <b>V</b> E S G C W M L Y E R E N Y Q G Q Q Y L L R R G E Y P D
	CRYGD	H P N L Q P Y L S R C N S A R <b>V</b> D S G C W M L Y E Q E N Y S G L Q Y F L R R G D Y A D
	CRYGS	C A D F H T Y L S R C N S I K <b>V</b> E G G T W A V Y E R E N F A G Y M Y I L P Q G E Y P E
		* * * * *

**Figure 2.**

Pedigree, DNA sequence and amino acid conservation associated with *CRYBA2* allele. **a.** Pedigree of family 4 with affected status and genotyping results indicated (+/-: heterozygous *CRYBA2* mutation present, -/-: *CRYBA2* mutation absent). The proband is indicated with an arrow; black symbol indicates congenital or juvenile cataract, gray symbol indicates adult-onset cataract. **b.** DNA sequence of *CRYBA2* showing heterozygous c.148G>A mutation (control sequence on bottom). **c.** Alignment of protein regions showing conservation of Val50 residue in *CRYBA2* proteins in various species as well as other beta- and gamma- crystallins.



**Figure 3.**

Expression of zebrafish *CRYBA2* orthologs, *cryba2a* and *cryba2b*, during embryonic development. Expression of *cryba2a* (**a, d**) and *cryba2b* (**b, c, e, f**) in whole mount embryos (**a, b, d, e**) and frontal sections of the head at the level of the eyes (**c, f**) in 22-hpf (**a–c**) and 48-hpf (**d–f**) embryos; lateral views (**a, b, d**) and a dorsal view (**e**) are shown. Please note strong and specific expression of both genes in the developing lens (l); on- optic nerve, r- retina.

Table 1

Summary of causative and likely causative mutations identified in this study (both known and novel cataract genes).

Patient #	Phenotype <sup>a</sup>	Race/ethnicity	Mutated Gene	Reference sequence	DNA effect	Protein effect	Cosegregation analysis (affected pedigrees)	Allele frequency (controls) <sup>b</sup>	PolyPhen/SIFT prediction	Previously reported
<b>Known cataract genes</b>										
<b>Patient 2</b>	Congenital cataract; glaucoma, <i>microcornea</i>	Caucasian (USA)	<i>CRYBB1</i>	NM_001887.3	c.286G>T	p.(Val96Phe)	Cosegregates: present in 6 affected, absent in 3 unaffected persons	0/12,971	Probably damaging/deleterious	No
<b>Patient 3</b>	Congenital cataract; <i>glaucoma</i>	Caucasian (USA)	<i>GJA8</i>	NM_005267.4	c.200A>G	p.(Asp67Gly)	Cosegregates: present in 2 affected persons	0/13,005	Probably damaging/deleterious	No
<b>Patient 6</b>	Congenital nuclear cataract, hyperopia, strabismus	Caucasian (Ashkenazi Jewish, Israel)	<i>CRYGD</i>	NM_006891.3	c.418C>T	p.(Arg140*)	Cosegregates: present in 3 affected, absent in 1 adult-onset and 2 unaffected persons	0/12,971	N/A (premature truncation)	Reviewed in Huang and He, 2010
<b>Patient 7</b>	Congenital cataract; <i>glaucoma</i>	Caucasian (Italy)	<i>CRYBB3</i>	NM_004076.3	c.581T>A	p.(Val194Glu)	Cosegregates with incomplete penetrance: present in 5 affected and 1 unaffected persons	0/12,999	Probably damaging/deleterious	No
<b>Patient 9</b>	Unilateral or bilateral congenital cataract	Caucasian (USA)	<i>CRYGC</i>	NM_020989.3	c.157_161 dupGC GGC	p.(Gln55 Valfs*50)	Cosegregates: present in 4 affected, absent in 5 unaffected persons	Not reported	N/A (frame shift)	Reviewed in Huang and He, 2010
<b>Patient 10</b>	Congenital cataract; <i>strabismus</i>	Caucasian (USA)	<i>MIP</i>	NM_012064.3	c.605G>A	p.(Trp202*)	Cosegregates: present in 3 affected, absent in 3 unaffected persons	0/12,178	N/A (premature truncation)	No
<b>Patient 11</b>	Congenital cataract; <i>pupil defects, glaucoma</i>	Caucasian (USA)	<i>GJA8</i>	NM_005267.4	c.226C>T	p.(Arg76Cys)	Cosegregates: present in 3 affected, absent in 2 unaffected persons	0/13,005	Probably damaging/deleterious	No
<b>Patient 15</b>	Congenital cataract	Unreported (USA)	<i>EYAL1</i>	NM_000503.4	c.121G>A	p.(Glu41Lys)	Not determined; no family members available for testing	0/12,949	Possibly damaging/tolerated	No
<b>Patient 18</b>	Congenital cataract; <i>microphthalmia/microcornea, corneal opacity, glaucoma</i>	Caucasian (USA)	<i>CRYGC</i>	NM_020989.3	c.417C>G	p.(Tyr139*)	Cosegregates: present in 2 affected, absent in 2 unaffected persons	0/12,971	N/A (premature truncation)	No
<b>Patient 22</b>	Congenital cataract, microcornea, macrocephaly; <i>coloboma, glaucoma</i>	Caucasian (Canada)	<i>CRYAA</i>	NM_000394.2	c.34C>T	p.(Arg12Cys)	Cosegregates: present in 3 affected persons	0/12,961	Probably damaging/deleterious	Reviewed in Huang and He, 2010
<b>Novel cataract genes</b>										
<b>Patient 4</b>	Congenital cataract; <i>myopia, glaucoma</i>	Hispanic (Uruguay)	<i>CRYBA2</i>	NM_057093.1	c.148G>A	p.(Val50Met)	Cosegregates with incomplete penetrance: present in 7 affected, 3 obligate carriers; absent in 3 unaffected persons	0/12,948	Probably damaging/deleterious	No

<sup>a</sup> variable features within the family are noted in italics; cataracts, unless specified, are bilateral;

<sup>b</sup> allele frequency provided for average number of chromosomes covered for the region according to the EYS database.