Exome sequencing identifies novel and recurrent mutations in GJA8 and CRYGD associated with inherited cataract

Donna S. Mackay
*Washington University School of Medicine in St. Louis*

Thomas M. Bennett
*Washington University School of Medicine in St. Louis*

Susan M. Culican
*Washington University School of Medicine in St. Louis*

Alan Shiels
*Washington University School of Medicine in St. Louis*

Follow this and additional works at: [https://digitalcommons.wustl.edu/open_access_pubs](https://digitalcommons.wustl.edu/open_access_pubs)

**Recommended Citation**
Mackay, Donna S.; Bennett, Thomas M.; Culican, Susan M.; and Shiels, Alan, "Exome sequencing identifies novel and recurrent mutations in GJA8 and CRYGD associated with inherited cataract." *Human Genomics*. 8, 19. (2014).
[https://digitalcommons.wustl.edu/open_access_pubs/3526](https://digitalcommons.wustl.edu/open_access_pubs/3526)

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact [vanam@wustl.edu](mailto:vanam@wustl.edu).
Exome sequencing identifies novel and recurrent mutations in GJA8 and CRYGD associated with inherited cataract

Donna S Mackay, Thomas M Bennett, Susan M Culican and Alan Shiels

Abstract

Background: Inherited cataract is a clinically important and genetically heterogeneous cause of visual impairment. Typically, it presents at an early age with or without other ocular/systemic signs and lacks clear phenotype-genotype correlation rendering both clinical classification and molecular diagnosis challenging. Here we have utilized trio-based whole exome sequencing to discover mutations in candidate genes underlying autosomal dominant cataract segregating in three nuclear families.

Results: In family A, we identified a recurrent heterozygous mutation in exon-2 of the gene encoding γD-crystallin (CRYGD; c.70C > A, p.Pro24Thr) that co-segregated with ‘coralliform’ lens opacities. Families B and C were found to harbor different novel variants in exon-2 of the gene coding for gap-junction protein α8 (GJA8; c.20T > C, p.Leu7Pro and c.293A > C, p.His98Pro). Each novel variant co-segregated with disease and was predicted in silico to have damaging effects on protein function.

Conclusions: Exome sequencing facilitates concurrent mutation-profiling of the burgeoning list of candidate genes for inherited cataract, and the results can provide enhanced clinical diagnosis and genetic counseling for affected families.

Keywords: Cataract, Exome sequencing, CRYGD, GJA8

Background

Hereditary forms of cataract constitute a clinically and genetically heterogeneous condition affecting the ocular lens [1-3]. Typically, inherited cataract has an early onset (<40 years) and most cases are diagnosed at birth (congenital), during infancy, or during childhood accounting for 10%–25% of all pediatric cataract cases [2]. Congenital and infantile forms of cataract are a clinically important cause of impaired visual development that accounts for 3%–39% of childhood blindness, worldwide [4]. Despite advances in surgical treatment, pediatric cataract poses a long-term risk of postoperative complications including secondary glaucoma, nystagmus, and retinal detachment [5-9].

Cataract can be inherited, either, as an isolated lens phenotype—usually with autosomal dominant transmission and full penetrance—or as part of a genetic/metabolic disorder (http://www.omim.org) involving additional ocular defects (e.g., anterior segment dysgenesis MIM107250) and/or systemic abnormalities (e.g., galactosemia MIM230400). Under slit-lamp examination, inherited cataract exhibits considerable inter- and intrafamilial phenotypic variation in location, size, shape, density, progression rate, and even color of the lens opacities [10]. Currently, genetic studies have identified over 39 genes and loci for inherited cataract, with or without other ocular signs [1,3]. These include gene coding for α-, β-, and γ-crystallins (e.g., CRYAA, CRYBB2, CRYGD), α-connexins (GJA3, GJA8) and other lens membrane or cytoskeleton proteins (e.g., MIP, BFSP2), several transcription factors (e.g., HSF4, PITX3), and an expanding group of functionally divergent genes (e.g., EPHA2, TDRD7, FYCO1). Since mutations in the same gene can cause morphologically different lens opacities and mutations in different genes can cause similar opacities, there is little genotype-phenotype correlation for inherited cataract.
rendering both clinical classification and molecular diagnosis challenging.

Traditionally, linkage analysis in extended pedigrees has been used to map cataract disease loci to specific chromosome regions and thereby limit the number of positional candidate genes that need to be conventionally sequenced in order to discover underlying mutations. However, the advent of next-generation (massively parallel) sequencing has facilitated the concurrent screening of multiple candidate genes in nuclear families and cases without a family history. Here, we have undertaken affected child-parent-trio-based whole-exome next-generation sequencing in order to identify mutations underlying autosomal dominant cataract in three nuclear families.

Results

Cataract families

We investigated three Caucasian-American pedigrees segregating cataract with autosomal dominant transmission in the absence of other ocular and/or systemic abnormalities (Figures 1A and 2A,D). A review of ophthalmic records indicated that bilateral cataract was diagnosed at birth (congenital) or during infancy in all three families with age-at-surgery ranging from 3 months to 1 year. In family A, the lens opacities appeared similar to those first described by Gunn in 1895 as resembling a piece of coral or coralliform [11]. No clinical images of lens opacities were available for family B or C, and none of the families had a sufficient number of meiotic events (≥10) to support independent linkage analysis. Instead, an affected child-parent plus spouse trio from each family was selected for whole-exome sequencing.

Candidate genes and exome sequences

We pre-selected 39 candidate genes for inherited cataract (Additional file 1) cited in the OMIM (http://www.omim.org), Cat-Map (http://cat-map.wustl.edu/), and iSyTE (http://bioinformatics.udel.edu/Research/iSyTE) databases [3,12]. The candidate list comprises genes known to be highly expressed in the lens including those coding for crystallins, connexins, and other lens membrane/cytoskeletal proteins, along with several more widely expressed genes that are associated with cataract and other limited eye/systemic conditions. Collectively, these candidate genes span over 111,000 bps of the genome and contain 300 exons located on chromosomes 1–13, 16, 17, 19–22, and X.

For all nine exome samples, over 98% of total paired-end reads were mapped to the reference genome (Additional file 2). Approximately 72%–84% of mapped reads were present in the captured exomes, and the average mean-mapped read-depth was 149.2X. With the exception of one sample in family C (C-I:1), >97% of each exome achieved a read-depth of ≥10X coverage, yielding a total of >38,900 single nucleotide polymorphisms (SNPs), of which >8,400 were non-synonymous and >1,400 were novel. For exome C-I:1, 80.61% reached ≥10X coverage yielding a total of 34,435 SNPs (7,639 non-synonymous and 1,331 novel). In addition, exome C-I:1 contained several more unexpected
regions of low coverage (gaps) than those detected in the other exomes (Additional file 2). However, the reduced coverage of exome C-I:1 did not compromise analysis of variants in the candidate genes of interest. Coverage of the 39 candidate genes exceeded a read-depth of >10X with three exceptions. The iron response element (IRE) of FTL is located in the 5′-UTR (untranslated region) and was not covered by capture probes. In addition, coverage of the single exons that code for FOXE3 and MAF was incomplete as previously reported [13]. We excluded mutations in all three missing gene regions by Sanger sequencing of an affected member of each family essentially as described [13,14]. Collectively, from the nine exomes sequenced, 112 variants were identified in 32 of the 39 candidate genes (Additional file 3). Of these variants, only five did not have genome reference sequence (rs) numbers and were potentially novel variants.

**Family A variants**

A review of the exome SNPs in family A with the list of candidate genes for cataract identified a total of 76 variants in 28 of 39 genes (Additional file 1 and Additional file 3). Of these, six variants (two coding/missense and four non-coding/synonymous) in five candidate genes were found in both affected relatives and not in the unaffected spouse. However, five of these variants associated with four candidate genes (SLC16A12, PAX6, CRYAB, GALK1) were excluded as disease-causing mutations as they have minor allele frequencies (MAFs) >0.01% (range 8.5%–52.4%) in Caucasians (Additional file 3). We note that the variant rs3740030 in SLC16A12 on chromosome 10 (chr10:91,222,287) has previously been associated with age-related cataract [15]. As rs3740030 was first thought to be a non-coding variant located in the 5′-UTR, the authors proposed a complex functional mechanism that...

![Figure 2](https://example.com/figure2.png)

**Figure 2** Mutation analysis of inherited cataract in family B and family C. (A) Pedigree of family B. The trio of individuals I:1, I:2, and I:1 was subject to exome sequencing. (B) Sanger sequence of GJA8 showing the heterozygous c.20 T > C, and p.Leu7Pro mutation found in affected individuals I:1, I:1, I:2, and I:3 but not in the affected spouse I:2 (lower trace). Horizontal bars indicate the codon reading frame. (C) Amino acid alignment of GJA8 showing high cross-species conservation of Leu7. (D) Pedigree showing family C. The trio of individuals I:1, I:2, and I:2 was subject to exome sequencing. (E) Sanger sequence of GJA8 showing the heterozygous c.293A > C and p.His98Pro mutation found in affected individuals I:1 and I:2 (upper trace) but not in the unaffected individuals I:2 and I:1 (lower trace). Horizontal bars indicate the codon reading-frame. (F) Amino acid alignment of GJA8 showing high cross-species conservation of His98.
involved modulation of translational efficiency. However, rs3740030 is now known to be located in exon-3 of SLC16A12 (c.49T > G) and was predicted to result in a non-conservative tryptophan-to-glycine substitution at codon 17 (p.Trp17Gly). While this variant was also predicted in silico to have a damaging effect on protein function (PolyPhen-2 score =0.997), it had a MAF value of 8.5% in Caucasians, suggesting that it is unlikely to be disease causing in family A. The remaining variant, rs28931605, occurred in exon-2 of CRYGD (c.70C > A) on chromosome 2 (chr2:208,989,018) and was predicted to result in the non-conservative substitution of proline-to-threonine at codon 24 (p.Pro24Thr) (Table 1). While this variant was predicted in silico to be tolerated, benign, or neutral with respect to protein function (Table 2), it has been previously associated with autosomal dominant cataract in multiple families (Additional file 4). The p.Pro24Thr variant has also been associated with autosomal dominant cataract in three nuclear families with inherited cataract [16]. Sanger sequencing of all five members of family B (Figure 2A,B) confirmed that the novel p.Leu7Pro variant in CRYGD associated with inherited cataract [16] and was predicted to result in a non-conservative substitution of histidine-to-proline at codon 98 (p.His98Pro) (Table 1). Histidine 98 is phylogenetically conserved across species (Figure 2F), and this variant was also predicted to have damaging effects on protein function using six mutation prediction programs (Table 2). Sanger sequencing of all four members of family C (Figure 2D,E) confirmed that the novel p.His98Pro variant in GJA8 co-segregated with cataract, consistent with it being the disease-causing mutation.

Family B variants

A review of the exome SNPs in family B with the candidate gene list, revealed a total of 73 variants in 22 of 39 genes (Additional file 1 and Additional file 3). Only 13 of these variants (12 non-coding or synonymous) associated with 7 of the candidate genes were found in both affected relatives and not in the unaffected spouse. All 12 non-coding or synonymous variants had MAF values >0.01% (range 0.4%–45.80%) and were effectively excluded as disease-causing mutations. The remaining variant was located in exon-2 of GJA8 (c.20T > C) on chromosome 1 (chr1:147,380,102) and was predicted to result in the substitution of leucine-to-proline at codon 7 (p.Leu7Pro) (Table 1). The p.Leu7Pro substitution represented a relatively conservative change with the non-polar, side-chain of leucine ([CH3]2-CH-CH2-) replaced by the unusual non-polar, side-ring of proline (−CH2-CH2−). However, Leu7 is phylogenetically conserved in GJA8 (Figure 2C), and the Pro7 substitution was predicted in silico to have probably damaging effects on protein function (Table 2). Sanger sequencing of all five members of family B (Figure 2A,B) confirmed that the novel p.Leu7Pro variant in GJA8 co-segregated with cataract further suggesting that it was the disease-causing mutation.

Family C variants

A review of the exome SNPs in family C using the candidate gene list yielded a total of 82 variants in 23 of the 39 genes (Additional file 1 and Additional file 3). However, only three of these variants associated with the candidate genes, WFS1, BFSP1, and GJA8, were present in both affected relatives but not in the unaffected spouse. The variants associated with WFS1 (rs734312) and BFSP1 (rs2281207) had MAF values of 54.69% and 25.74%, respectively, and were excluded as causative mutations. The remaining variant occurred in exon-2 of GJA8 (c.293A > C) on chromosome 1 (chr1:147,380,375) and was predicted to result in a non-conservative substitution of histidine-to-proline at codon 98 (p.His98Pro) (Table 1). Histidine 98 is phylogenetically conserved across species (Figure 2F), and this variant was also predicted to have damaging effects on protein function using six mutation prediction programs (Table 2). Sanger sequencing of all four members of family C (Figure 2D,E) confirmed that the novel p.His98Pro variant in GJA8 co-segregated with cataract, consistent with it being the disease-causing mutation.

Discussion

Several recent studies have employed exome sequencing of index patients or probands in multiple families in order to discover mutations in candidate genes underlying autosomal dominant and recessive forms of cataract [13,17–19]. In this study, we have used trio-based exome sequencing to uncover a recurrent missense mutation in CRYGD (p.Pro24Thr) and two novel missense mutations in GJA8 (p.Leu7Pro, p.His98Pro) associated with autosomal dominant cataract in three nuclear families. Child–parent trios offer the initial benefit of co-segregation testing during exome variant analysis, but this advantage may be offset in larger cohorts of families by the additional sequencing costs. The p.Pro24Thr

Table 1 Summary of mutations detected by exome sequencing of trios from families A, B, and C

| Pedigree (numbers affected) | Physical location (chr:bp) | Candidate gene (ID) | Exon | cDNA variant | Protein variant | Allele frequency (EVS)* | Status |
|-----------------------------|---------------------------|---------------------|------|--------------|------------------|------------------------|--------|
| A (3)                       | chr2:208,989,018          | CRYGD (1412)        | 2    | c.70C > A    | p.Pro24Thr       | 0/8,600                | Recurrent (Additional file 5) |
| B (4)                       | chr1:147,380,102          | GJA8 (2703)         | 2    | c.20T > C    | p.Leu7Pro        | 0/8,600                | Novel  |
| C (2)                       | chr1:147,380,375          | GJA8 (2703)         | 2    | c.293A > C   | p.His98Pro       | 0/8,600                | Novel  |

*Allele frequencies for European Americans listed on the Exome Variant Server.
substitution in *CRYGD* has now been identified in some 14 different families, mostly segregating coralliform cataract that affects more than 133 individuals with varied ethnic backgrounds and constitutes the most recurrent missense mutation in a crystallin gene to be associated with inherited cataract (Additional file 4). The novel mutations found in *GJA8* increase the mutation spectrum of this connexin gene to at least 32 different mutations segregating in 38 families making it one of the most common non-crystallin genes to be associated with inherited cataract in humans (Additional file 5).

*CRYGD* (MIM: 123690) consists of three exons and encodes γD-crystallin—a hydrophylic protein of 174 amino acids that is characterized by two βγ-crystallin domains each formed by two repeat Greek-key motifs of approximately 40 residues. *CRYGD* is expressed at high concentrations in fiber cells of mammalian lenses and plays an important structural role in establishing lens transparency and gradient refractive index [20]. Proline at position 24 is located within the first Greek-key motif of human *CRYGD* but is not well conserved across species (replaced by serine in the mouse and threonine in the zebrafish). Consequently, *in silico* analysis predicted that the Pro24Thr substitution was benign (Table 2). Further, NMR-spectroscopy and X-ray crystallography have indicated that the Pro24 and Thr24 proteins are structurally similar overall [21,22]. However, the Thr24 mutant exhibits local conformational and dynamic differences that may initiate aggregation or polymerization and *in vitro* experiments have shown that the Thr24 protein exhibits reduced solubility—a property that is likely to trigger cataract formation [23-25].

*GJA8* (MIM: 600897) comprises two exons with exon-2 coding for the entire 433 amino acid residues of gap-junction protein α8 or connexin 50. *GJA8* contains four transmembrane domains that are joined by two extracellular loops and one cytoplasmic loop and flanked by cytoplasmic N- and C-termini. By forming hexamers, or hemi-channels, that can dock between adjacent cells to create gap-junction channels, *GJA8* plays an important role in lens intercellular communication [26]. Of the 32 known coding mutations in *GJA8*, 30 result in missense substitutions that, with one exception, are associated with autosomal dominant cataract, and the remaining two are frameshift mutations associated with autosomal recessive cataract (Additional file 5). Most of the missense substitutions are located within the N-terminal half of the protein, which also contains the conserved connexin domain (pfam00029; amino acids 3–109). The novel p.Leu7Pro substitution found in family B is the first to be located at the cytosolic N-terminal end of human *GJA8*. Support for its pathogenicity in humans is provided by the SHR-Dca rat strain, which inherits semi-dominant cataract [27]. Heterozygous (+/Dca) mutants develop nuclear pulverulent opacities and smaller eyes than wild-type, while homozygotes (Dca/Dca) present with severe microphthalmia and a hypoplastic lens. The underlying mutation has been identified as a missense mutation in *GJA8* (c.20T > A) that is predicted to result in a non-conservative p.Leu7Gln substitution. Both the rat p.Leu7Gln and

| Prediction program | Prediction | PMUT | PANTHER | MutPred |
|--------------------|------------|------|---------|---------|
| **CRYGD** | c.70C > A | p.Pro24Thr | 0.10 | 0.0936 |
| SIFT | Value (<0.05) | 0.10 | 0.0936 |
| | Prediction | Tolerated | Neutral |
| Polyphen-2 | Score | 0.084 | 0.0936 |
| | Prediction | Benign | Neutral |
| PMUT | NN output | 0.0936 | 0.0936 |
| | Reliability | 8 | 8 |
| | Prediction | Neutral | Neutral |
| PANTHER | subPSEC (<–3) | –2.35974 | –2.35974 |
| | P deleterious (>0.5) | 0.34519 | 0.34519 |
| PON-P2 | Probability for pathogenicity | 0.479 | 0.479 |
| | Standard error | 0.332 | 0.332 |
| | Prediction | Unknown | Unknown |
| MutPred | Probability of deleterious mutation | 0.840 | 0.840 |
| | Molecular mechanisms disrupted | Loss of stability (p =0.0067) | Loss of stability (p =0.0067) |

| **GJA8** | c.20T > C | p.Leu7Pro | 0.918 |
| SIFT | Value (<0.05) | 0.0936 | 0.0936 |
| | Prediction | Tolerated | Neutral |
| Polyphen-2 | Score | 0.991 | 0.918 |
| | Prediction | Probably damaging | Pathological |
| PMUT | NN output | 0.8749 | 0.8749 |
| | Reliability | 7 | 7 |
| | Prediction | Pathological | Pathological |
| PANTHER | subPSEC (<–3) | –4.3388 | –4.3388 |
| | P deleterious (>0.5) | 0.79148 | 0.79148 |
| PON-P2 | Probability for pathogenicity | 0.947 | 0.947 |
| | Standard error | 0.050 | 0.050 |
| | Prediction | Pathogenic | Pathogenic |
| MutPred | Probability of deleterious mutation | 0.889 | 0.889 |
| | Molecular mechanisms disrupted | Gain of disorder (p =0.0099) | Gain of disorder (p =0.0099) |

| **GJA8** | c.293A > C | p.His98Pro | 0.000 |
| SIFT | Value (<0.05) | 0.000 | 0.000 |
| | Prediction | Not tolerated | Not tolerated |
| Polyphen-2 | Score | 1.000 | 1.000 |
| | Prediction | Probably damaging | Probably damaging |
| PMUT | NN output | 0.7822 | 0.7822 |
| | Reliability | 5 | 5 |
| | Prediction | Pathological | Pathological |
| PANTHER | subPSEC (<–3) | –3.98807 | –3.98807 |
| | P deleterious (>0.5) | 0.72871 | 0.72871 |
| PON-P2 | Probability for pathogenicity | 0.444 | 0.444 |
| | Standard error | 0.090 | 0.090 |
| | Prediction | Unknown | Unknown |
| MutPred | Probability of deleterious mutation | 0.889 | 0.889 |
| | Molecular mechanisms disrupted | Loss of stability (p =0.0067) | Loss of stability (p =0.0067) |

Table 2 *In silico* predictions of functional effects for the three mutations identified in this study
human p.Leu7Pro mutations result in the substitution of a highly conserved leucine residue with uncharged amino acids, suggesting that they may exert similar deleterious effects on G/A8 function.

The novel p.His98Pro mutation identified in family C, is located near the junction of the second transmembrane domain with the cytoplasmic loop of G/A8. Four other mutations, p.Val79Leu, p.Pro88Ser, p.Pro88Gln, and p.Pro88Thr, have previously been localized to the second transmembrane domain. Based on the non-conservative nature of the p.His98Pro substitution, we speculate that this mutant will also fail to reach the plasma membrane and form functional gap-junction channels.

Conclusions
Exome sequencing provides a rational approach to concurrently screen over 39 candidate genes for inherited cataract in nuclear families or even sporadic cases. In addition, exome sequencing may enable the discovery of novel genes underlying inherited cataract and, potentially, genes associated with age-related cataract. However, considerable supporting evidence (e.g., additional mutations, functional expression in vitro, and/or an animal model) will be required to verify disease causation. In a clinical setting, results from exome sequencing are unlikely to be ‘clinically actionable’ with respect to surgical treatment and subsequent management of inherited cataract. However, such data can contribute to a gene-centric clinical classification of inherited cataract and provide enhanced diagnosis and genetic counseling for affected families.

Methods
Ethics statement
Ethical approval for this study was obtained from the Washington University Human Research Protection Office (HRPO), and written informed consent was provided by all participants prior to enrollment in accordance with the tenets of the Declaration of Helsinki and Health Insurance Portability and Accountability Act (HIPAA) regulations.

Family participants
Three Caucasian-American pedigrees segregating autosomal dominant cataract were ascertained through ophthalmic records in the Department of Ophthalmology and Visual Sciences at Washington University School of Medicine. Blood samples were obtained from available family members including a spouse (Figures 1 and 2). Leukocyte genomic DNA was purified using the Gentra Puregene Blood kit (Qiagen, Valencia, CA) and quantified by absorbance at 260 nm (NanoDrop 2000, Wilmington, DE).

Exome sequencing
Whole exome capture was achieved using the SureSelect Human All Exon V5 (50.4 Mb) Kit, according to manufacturer’s instructions (Agilent Technologies). Genomic DNA (3 μg) was fragmented (150–200 bp) by acoustic shearing, ligated to adapter primers, and PCR-amplified. Following denaturation (95°C, 5 min), amplified DNA-fragment libraries (~500 ng) were hybridized in a solution under high stringency (65°C, 24 h) with biotinylated RNA capture probes (~120 bp). Resulting DNA/RNA hybrids were recovered by streptavidin-coated magnetic bead separation (Dynal, Invitrogen, Carlsbad, CA). Captured DNA was eluted (NaOH) and then subjected to solid phase (flow-cell) next-generation (massively parallel) sequencing on a HiSeq2000 System (Illumina, San Diego, CA) using the Illumina Multiplexing Sample Preparation Oligo-nucleotide Kit and the HiSeq 2000 Paired-End Cluster Generation Kit according to the manufacturer’s instructions. Briefly, hybrid-capture libraries were amplified to add indexing (identifying) tags and sequencing primers then subjected to paired-end (2 × 101 bp read length), multiplex sequencing-by-synthesis using fluorescent, cyclic reversible (3′-blocked) terminators. A pool of three exome samples (representing a family trio) was sequenced in a single lane of the sequencer’s flow-cell.

Exome variant analysis
Raw sequence data was aligned to the human reference genome (build hg19) by NovoalignMPI (www.novocraft.com), and sequence variants called using the Sequence Alignment/Map format (SAMtools) and Picard programs (http://samtools.sourceforge.net/) and further annotated using SeattleSeq (http://snp.gs.washington.edu/SeattleSeqAnnotation138/). Target coverage and read-depth were reviewed by the Integrated Genomics Viewer (IGV; http://www.broadinstitute.org/igv/). Variants were filtered using the Ingenuity variant analysis website (IVA http://ingenuity.com) or the gnomAD project pipeline (http://gnome.tchlab.org/) [29]. Identified variants in the pre-selected candidate genes (Additional file 1) were then reviewed for presence/absence and frequency in various websites including dbSNP (http://www.ncbi.nlm.nih.gov/snp/), 1000 genomes (http://www.1000genomes.org/), and the Exome Variant Server database (http://evs.gs.washington.edu/EVS/). The predicted effect on protein function was analyzed using the SIFT (http://sift.jcvi.org), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), PMUT (http://mmb2.pcb.ub.es:8080/PMut/), and/or an animal model will be required to verify disease causation. In a clinical setting, results from exome sequencing are unlikely to be ‘clinically actionable’ with respect to surgical treatment and subsequent management of inherited cataract. However, such data can contribute to a gene-centric clinical classification of inherited cataract and provide enhanced diagnosis and genetic counseling for affected families.
PON-P2 (http://structure.bmc.lu.se/PON-P2/), PANTHER (http://www.pantherdb.org/tools/csnpScoreForm.jsp), and MutPred (http://mutpred.mutdb.org/) in silico mutation prediction programs [30-34].

Sanger sequencing
Genomic DNA (2.5 ng/μl, 10 μl reactions) was amplified (35 cycles) in a GeneAmp 9700 thermal cycler using Top Taq mastermix kit (Qiagen) and 20 pmol of gene-specific primers (Additional file 6). Resulting PCR amplicons were enzyme-purified with ExoSAP-IT (USB Corporation, Cleveland, OH). The purified amplicons were direct cycle-sequenced in both directions with BigDye Terminator Ready Reaction Mix (v3.1)(Applied Biosystems, Grand Island, NY) containing M13 forward or reverse sequencing primers (Additional file 6). Resulting PCR amplicons were sequenced in both directions with BigDye Terminator Ready Reaction Mix (v3.1)(Applied Biosystems, Grand Island, NY) containing M13 forward or reverse sequencing primers (Additional file 6). Resulting PCR amplicons were then ethanol precipitated and detected by capillary electrophoresis on a 3130xl Genetic Analyzer running Sequence Analysis (v.6.0) software (Applied Biosystems) and Chromas (v2.23) software (Technelysium, Tewantin, Queensland, Australia).

Additional files

Additional file 1: Table S1. Candidate genes for inherited cataract evaluated in this study.

Additional file 2: Table S2. Sample metrics for exome sequencing of trios from families A, B, and C.

Additional file 3: Table S3. Exome sequencing variants found in candidate genes for inherited cataract (Additional file 1, Table S1) in family trios A, B, and C.

Additional file 4: Table S4. Recurrence of the p.Pro24Thr mutation in GJA8 in a control trio.

Additional file 5: Table S5. Mutation profile of GJA8.

Additional file 6: Table S6. PCR primers used for Sanger sequencing of mutations found in GJA8 and GJA2. Each primer used was tagged with ‘tagagagacagctagat’ and the reverse primers with ‘caggaaacagctatgacc’.

Abbreviations
OMIM: Online Mendelian Inheritance in Man; MAF: Minor allele frequency; SHR-Dca: Spontaneously hypertensive rat-Dominant cataract; NMR: Nuclear magnetic resonance.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
DSM and TMB were involved in acquisition and analysis of exome sequencing data and bioinformatics analyses. SMC coordinated ascertainment and recruitment of patients and was involved in acquisition and analysis of clinical data. DSM and AS conceived the study, participated in its design and coordination, and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements
We thank the families for their participation in this study and the Genome Technology Access Center (GTAC) at Washington University School of Medicine for help with genomic analysis. GTAC is partially supported by National Institutes of Health (NIH) grants P30 CA91842 and U11 TR000448. This work was supported by NIH grants EY012284 (to A.S.) and EY02687 (Core Grant for Vision Research) and an unrestricted grant to the Department of Ophthalmology and Visual Sciences from Research to Prevent Blindness (RPB).

Received: 27 August 2014 Accepted: 23 October 2014
Published online: 18 November 2014

References
1. Shiels A, Bennett T, Hejtmancik J: Cat-Map: putting cataract on the map. Mol Vis 2010, 16:2007–2015.
2. Trumler A: Evaluation of pediatric cataracts and systemic disorders. Curr Opin Ophthalmol 2011, 22:365–379.
3. Shiels A, Hejtmancik JF: Genetics of human cataract. Clin Genet 2013, 84:120–127.
4. Mickler C, Boden J, Trivedi RH, Wilson ME: Pediatric cataract. Pediatr Ann 2011, 40:63–87.
5. Matafsi A, Hadjich AB, Kkokkali S, Rabiah PK, Birch E, Stager DRJ, Cheong-Leen R, Singh V, Egbert J, Aistle WF, Lambert SR, Amtsh B, Kan AO, Grigg J, Avantidou M, Dimitriadou SA, Nischal KK: Postoperative glaucoma following infantile cataract surgery: an individual patient data meta-analysis. JAMA Ophthalmol 2014, doi:10.1001/jamaophthalmol.2014.1042. Published online June 12, 2014.
6. Lambert SR, Purohit A, Superal HM, Lynn MJ, Beck AD: Long-term risk of glaucoma after congenital cataract surgery. Am J Ophthalmol 2013, 155:356–361. e352.
7. Ruddle JB, Staffieri SE, Crowston JG, Shenwin JC, Mackey DA: Incidence and predictors of glaucoma following surgery for congenital cataract in the first year of life in Victoria, Australia. Clin Experiment Ophthalmol 2013, 41:653–661.
8. Young MP, Heidary G, VanderVeen DK: Relationship between the timing of cataract surgery and development of nystagmus in patients with bilateral infantile cataracts. J AAP 2012, 165:544–557.
9. Haagbard B, Anderson EW, Oudin A, Poulsen G, Wohlfarth J, La Cour M, Melbye M: Risk of retinal detachment after pediatric cataract surgery. Invest Ophthalmol Vis Sci 2014, 55:2947–2951.
10. Araya L, Taylor D, Russell-Eggitt I, McIvor L, Nischal KK, Lengyel D: The morphology and natural history of childhood cataracts. Surv Ophthalmol 2003, 48:125–144.
11. Gunn RH: Peculiar catarilliform cataract with crystals in the lens. Trans Ophthalmol Soc UK 1989, 109:219.
12. Lachke SA, Ho JW, Iyukov GJ, O’Connell DJ, Aboukhail A, Bulyk ML, Park PJ, Maas RL: iSyTE: integrated Systems Tool for Eye gene discovery. Invest Ophthalmol Vis Sci 2012, 53:1617–1627.
13. Reis LM, Tyler RC, Muheisen S, Raggio V, Salvati L, Han DP, Costakos D, Yonath H, Hall S, Power P, Semina EV: Whole exome sequencing in dominant cataract identifies a new causative factor, CRYBA2, and a variety of novel alleles in known genes. Hum Genet 2013, 132:161–170.
14. Bennett TM, Maraini G, Jin C, Sun W, Hejtmancik JF, Shiels A: Noncoding variation of the gene for ferritin light chain in hereditary and age-related cataract. Mol Vis 2013, 19:835–844.
15. Zuercher J, Neidhardt J, Magyar I, Labs S, Moore AT, Tanner FC, Waseen N, Schorideret DF, Munier FL, Bhattacharya S, Berger W, Kloeckner-Gruissim B: Alterations of the 5′ untranslated region of SLC16A12 lead to age-related cataract. Invest Ophthalmol Vis Sci 2010, 51:3354–3361.
16. den Dunnen JT, Antonakis SE: Nomenclature for the description of human sequence variations. Hum Genet 2001, 109:121–124.
17. Aldahmesh MA, Khan AO, Mohamed JR, Hijazi H, Al-Owain M, Alowaid A, Alkuwari FS: Genomic analysis of pediatric cataract in Saudi Arabia reveals novel candidate disease genes. Genet Med 2012, 14:955–962.
18. Sun W, Xiao X, Li S, Guo X, Zhang Q: Exome sequencing of 18 Chinese families with congenital cataracts: a new sight of the NHS gene. PLoS One 2014, 9(12):e100545.
19. Prokudin I, Simons C, Grigg JR, Storen R, Kumar V, Phua ZY, Smith J, Flaherty M, Davila S, Jamieson RV: Exome sequencing in developmental eye disease leads to identification of causal variants in GJA8, CYR6, PXAS and CYP1B1. Eur J Hum Genet 2014, 22:907–915.
20. Slingsby C, Wistow GJ, Clark AR: The structure of the cataract-causing P23T mutant of human gammaD-crystallin exhibits distinctive local conformational and dynamic changes. Biochemistry 2009, 48:2597–2609.
22. Ji F, Koharudin LM, Jung J, Gronenborn AM: Crystal structure of the cataract-causing P23T gammaD-crystallin mutant. Proteins 2013, 81:1493–1498.

23. Mackay D, Andley U, Shiels A: A missense mutation in the gammaD crystallin gene (CRYGD) associated with autosomal dominant “coral-like” cataract linked to chromosome 2q. Mol Vis 2004, 10:155–162.

24. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace BA, Singsby C: The P23T cataract mutation causes loss of solubility of folded yD-crystallin. J Mol Biol 2004, 343:435–444.

25. Pande A, Ghosh KS, Banerjee PR, Pande J: Increase in surface hydrophobicity of the cataract-associated P23T mutant of human gammaD-crystallin is responsible for its dramatically lower, retrograde solubility. Biochemistry 2010, 49:6122–6129.

26. Boyer E, Elahara L, Berthoud V: Connexin mutants and cataracts. Front Pharmacol 2013, 4:43.

27. Liška F, Chylíková B, Martinek J, Kren V: Microphthalmia and cataract in rats with a novel point mutation in connexin 50–L7Q. Mol Vis 2008, 7:628–632.

28. Rubinos C, Villone K, Mhaske P, White TW, Srinivas M: Functional effects of Cx50 mutations associated with congenital cataracts. Am J Physiol Cell Physiol 2014, 306:C212–C220.

29. Lee H, Lee K, Hsing M, Choe Y, Park JH, Kim SH, Bohn JM, Neu MB, Hwang KB, Green RC, Kohane IS, Kong SW: Prioritizing disease-linked variants, genes, and pathways with an interactive whole-genome analysis pipeline. Hum Mut 2014, 35:537–547.

30. Ferrer-Costa C, Gelpi JL, Zarnakola L, Parraga I, de la Cruz X, Orozco M: PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics 2005, 21:3176–3178.

31. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A: PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003, 13:2129–2131.

32. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: A method and server for predicting damaging missense mutations. Nat Methods 2010, 7:248–249.

33. Li B, Krishnan VG, Mort ME, Xin F, Karnati KK, Cooper DN, Mooney SD, Radivojac P: Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 2009, 25:2744–2750.

34. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC: SIFT web server: predicting effects of amino acid substitutions on proteins. Nucl Acids Res 2012, 40:W452–W457.

doi:10.1186/s40246-014-0019-6

Cite this article as: Mackay et al.: Exome sequencing identifies novel and recurrent mutations in GJA8 and CRYGD associated with inherited cataract. Human Genomics 2014 8:19.