RESEARCH ARTICLE

Deletion at the *GCNT2* Locus Causes Autosomal Recessive Congenital Cataracts

Bushra Irum[1,2], Shahid Y. Khan[1], Muhammad Ali[1], Muhammad Daud[1], Firoz Kabir[1], Bushra Rauf[1,2], Fareeha Fatima[2], Hira Iqbal[2], Arif O. Khan[3], Saif Al Obaisi[3], Muhammad Asif Naeem[2], Idrees A. Nasir[2], Shaheen N. Khan[2], Tayyab Husnain[2], Sheikh Riazuddin[2,4,5], Javed Akram[4,5], Allen O. Eghrari[1], S. Amer Riazuddin[1,6] *

1 The Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, MD, United States of America, 2 National Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan, 3 King Khaled Eye Specialist Hospital, Riyadh, Saudi Arabia, 4 Allama Iqbal Medical College, University of Health Sciences, Lahore, Pakistan, 5 National Centre for Genetic Diseases, Shaheed Zulfiqar Ali Bhutto Medical University, Islamabad, Pakistan, 6 McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States of America

* riazuddin@jhmi.edu

OPEN ACCESS

Abstract

Purpose

The aim of this study is to identify the molecular basis of autosomal recessive congenital cataracts (arCC) in a large consanguineous pedigree.

Methods

All participating individuals underwent a detailed ophthalmic examination. Each patient's medical history, particularly of cataracts and other ocular abnormalities, was compiled from available medical records and interviews with family elders. Blood samples were donated by all participating family members and used to extract genomic DNA. Genetic analysis was performed to rule out linkage to known arCC loci and genes. Whole-exome sequencing libraries were prepared and paired-end sequenced. A large deletion was found that segregated with arCC in the family, and chromosome walking was conducted to estimate the proximal and distal boundaries of the deletion mutation.

Results

Exclusion and linkage analysis suggested linkage to a region of chromosome 6p24 harboring *GCNT2* (glucosaminyl (N-acetyl) transferase 2) with a two-point logarithm of odds score of 5.78. PCR amplifications of the coding exons of *GCNT2* failed in individuals with arCC, and whole-exome data analysis revealed a large deletion on chromosome 6p in the region harboring *GCNT2*. Chromosomal walking using multiple primer pairs delineated the extent of the deletion to approximately 190 kb. Interestingly, a failure to amplify a junctional fragment of the deletion break strongly suggests an insertion in addition to the large deletion.

Conclusion

Here, we report a novel insertion/deletion mutation at the *GCNT2* locus that is responsible for congenital cataracts in a large consanguineous family.

Introduction

The ocular lens focuses light that passes through it onto the retina, where it is detected by photoreceptors and transformed into visual signals [1]. Cataract is an opacity or cloudiness caused by damage to the precise cellular structure of the lens, the precision of which is crucial to maintaining the transparency of the lens [2,3]. Light-blocking opacities compromise the function of the lens and reduce vision, often severely, unless the cataractous lens is removed [2,3]. Congenital cataracts are the primary cause of childhood blindness [2]. They occur in 1–6 cases per 10,000 live births; however, the incidence is much higher (5–15 cases per 10,000 live births) in developing countries [3,4].

Inherited cataracts comprise a significant fraction of the global burden of cataractogenesis [5]. Congenital cataracts may occur as an isolated anomaly only affecting the ocular lens or in conjunction with a developmental dystrophy of the anterior segment, such as microphthalmia or Peters anomaly [6]. Additionally, congenital cataracts have been associated with genetic multisystem disorders, such as Lowe syndrome and Nance-Horan syndrome [6]. More than 40 loci linked to human congenital cataracts have been mapped on different chromosomes [5]. Even phenotypically identical cataracts can be caused by mutations at different genetic loci (http://cat-map.wustl.edu/) and may exhibit different inheritance patterns [7].

Autosomal recessive congenital cataracts (arCC) have been associated with loci and genes on several chromosomes (1p, 1q, 3p, 3q, 6p, 7q, 8p, 9q, 11q, 16q, 17q, 19q, 20p, 21q, and 22q) [8–24]. Pathogenic mutations have been indentified in EPH receptor A2 (*EPHA2*), connexin50 (*GJA8*), FYVE and coiled-coil domain containing 1 (*FYCO1*), glucosaminyl (N-acetyl) transferase 2 (*GCNT2*), acylglycerol kinase (*AGK*), tudor domain containing 7 (*TDRD7*), crystallin alpha B (*CRYAB*), heat-shock transcription factor 4 (*HSF4*), galactokinase 1 (*GALK1*), lens intrinsic membrane protein 2 (*LIM2*), beaded filament structural protein 1 (*BFSP1*), crystallin alpha A (*CRYAA*), lanosterol synthase (*LSS*), crystallin beta B1 (*CRYBB1*), and crystallin beta B3 (*CRYBB3*) [9,10,13,17–19,21–29].

GCNT2, the gene implicated in this study, encodes the I-branching enzyme, a beta-1,6-N-acetylglucosaminyltransferase that converts fetal i antigen in erythrocytes to the adult I antigen [13]. The i antigen is a linear poly-N-acetyllactosamine chain, whereas the I-antigen structure contains branched poly-N-acetyllactosaminoglycans [30,31]. First discovered in human red blood cells with cold-agglutinating autoantibodies [32,33], the I and i antigens are carbohydrate structures attached to glycolipids and glycoproteins on the cell surfaces of various tissues and body fluids [31,34]. Normally red blood cells predominantly express the I antigen on the cellular surface; however, in patients with the adult i phenotype, red blood cells express low levels of I antigen and high levels of the i antigen [35].

Yamaguchi and colleagues were among the first to report the association of the adult i phenotype with congenital cataracts [36]. To date, multiple mutations in *GCNT2* have been identified in patients with congenital cataracts. These include both missense and nonsense mutations in *GCNT2* [35,37–39]. Moreover, deletions of coding exons of *GCNT2* have been reported [35,40]. Here, we report the causal mutation in a large, consanguineous familial case of congenital cataracts. With linkage analysis, we localized the disease interval to a region on

chromosome 6p harboring *GCNT2*, and failure to amplify the gene in affected individuals suggested an insertion/deletion at the locus. Whole-exome sequencing coupled with chromosome walking confirmed a deletion of approximately 190 kb, including *GCNT2*.

Materials and Methods

Patient Recruitment and Clinical Evaluation

Consanguineous Pakistani families with cataracts were recruited to participate in a collaborative study to understand the genetic aspects of arCC. Institutional Review Board approval was granted by the National Eye Institute (Bethesda, MD), the Johns Hopkins University School of Medicine (Baltimore, MD), and the National Centre of Excellence in Molecular Biology (Lahore, Pakistan). All participating subjects gave informed written consent consistent with the tenets of the Declaration of Helsinki.

A detailed medical history was compiled for each participant by interviewing family members and reviewing available medical records. Ophthalmic examinations, including slit-lamp microscopy, were performed at the Layton Rahmatulla Benevolent Trust Hospital in Lahore, Pakistan. All participating members donated approximately 10 ml of blood, drawn into 50 ml Sterilin Falcon tubes with 20 mM EDTA. Genomic DNA was extracted as previously described [41,42].

Exclusion Analysis

All documented arCC loci and genes were excluded by genotyping with 70 polymorphic short tandem repeat markers. Amplification reactions for the exclusion analysis was performed as described previously [41,42]. The amplified products from each DNA sample were mixed with a loading cocktail containing 400HD size standards (Applied Biosystems, Foster City, CA), and resolved on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Genotypes were assigned with GeneMapper software (Applied Biosystems).

Linkage Analysis

Two-point linkage analysis was performed using the FASTLINK version of MLINK from the LINKAGE Program Package (provided in the public domain by the Human Genome Mapping Project Resources Centre, Cambridge, UK) [43,44]. Maximum logarithm of odds (LOD) scores were calculated using PLINK (Shaun Purcell, Boston, MA). arCC was analyzed as a fully penetrant trait with an affected allele frequency of 0.001. The marker order and distances between the markers were calculated from the chromosome 6 sequence maps published by the NCBI (National Center for Biotechnology Information, Bethesda, MD).

Sanger Sequencing for Mutation Screening

Primer pairs for individual exons were designed using the Primer3 program. The primer sequences and amplification conditions are available upon request. The PCR amplifications were performed in 10 μl reactions with 20 ng genomic DNA as described previously [41,42]. The PCR primers for each exon were used for bidirectional sequencing with the BigDye Terminator Ready Reaction Mix, in accordance with the manufacturer's instructions (Thermo Fisher Scientific, Waltham, MA).

Sequencing products were dissolved in 10 μl of Formamide (Applied Biosystems) and resolved on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Sequencing results were assembled with ABI PRISM sequencing analysis software, version 3.7, and analyzed with SeqScape software (Applied Biosystems).

Chromosome Walking

A total of 23 primer pairs spanning the *GCNT2* locus were designed using Primer3 software (Table 1). PCR amplifications were performed in 10 µl reactions with 20 ng genomic DNA, 1 µl of 10 µM of forward and reverse primers, 1 µl of 10x PCR buffer (100 mM Tris HCl (pH 8.4), 400 mM NaCl, 15 mM $MgCl_2$, and 2.5 mM spermidine), 250 µM dNTP mix, 700 mM dimethyl sulfoxide (DMSO), 500 mM betaine, and 0.2 U Taq DNA polymerase. PCR amplification consisted of a denaturation step at 95˚C for 5 minutes followed by a two-step touchdown procedure. The first step of 10 cycles consisted of denaturation at 95˚C for 30 s, followed by primer-set-specific annealing for 30 s, and elongation at 72˚C for 45 s. The annealing temperature (provided in Table 1) decreased by 1˚C per cycle. The second step of 30 cycles consisted of denaturation at 95˚C for 30 s, followed by annealing (10˚C below annealing temperature used in the first step) for 30 s and elongation at 72˚C for 45 s. The last step was a final elongation at 72˚C for 5 minutes. The PCR products were separated on a 1.5% agarose gel and visualized on a UV transilluminator to determine the amplification.

Whole-Exome Library Preparation and Sequencing

Whole-exome sequencing libraries were prepared using the Nextera Rapid Capture Expanded Exome kit in accordance with the manufacturer's protocol (Illumina, San Diego, CA). Genomic DNA was quantified with a Fluorometer (Qubit 2.0; Invitrogen, Carlsbad, CA). The enriched libraries were quantitated with an Agilent 2100 Bioanalyzer (Palo Alto, CA) and pooled together in equal concentrations. The pooled, bar-coded exome libraries were clustered using the TruSeq v3 Cluster Kit at a 13 pM concentration, and paired-end sequenced (2 × 100 bp) using the TruSeq SBS Kit v3 on a HiSeq 2000 sequencing system (Illumina).

Sequence Alignment and Variant Calling

Bioinformatic analysis of exome data was performed with Lasergene Genomics Suite (DNASTAR, Madison, WI). The raw reads were mapped to the human genome (GRCh38) using DNASTAR's proprietary assembly software (SeqMan NGen 12) with default parameters. The mapped reads were subjected to ArrayStar, to filter the variants based on genotype, sequencing depth, conservation, and read-quality parameters. Additional filtering steps were applied to remove any variants with a minor allele frequency greater than 0.1% in public databases, including dbSNP, 1000 Genomes, exome variant servers (NHLBI server and ExAC browser), or in our in-house exome datasets of ethnically matched control samples and individuals from unrelated familial cases without cataracts.

Real-time Expression Analysis

The use of mice in this study was approved by the Johns Hopkins Animal Care and Use Committee (ACUC), and all experiments were performed by a protocol approved by the Johns Hopkins ACUC. Mouse lenses were obtained at different developmental stages, including embryonic day 15 (E15), day 18 (E18), at birth (P0), postnatal day 3 (P3), day 6 (P6), day 9 (P9), day 12 (P12), day 14 (P14), day 21 (P21), day 28 (P28), day 42 (P42), and day 56 (P56). Mouse lenses were obtained as described previously [41,42]. Total RNA from the lenses was isolated with TRIzol (Invitrogen) according to the manufacturer's instructions. The quality and quantity of the total RNA were determined on a NanoDrop Lite Spectrophotometer (Thermo Fisher Scientific). First-strand cDNA synthesis was completed with the SuperScript III First-Strand Synthesis Kit according to the manufacturer's instructions (Thermo Fisher Scientific). Quantitative real-time PCR was performed on a StepOne Real-Time PCR System

Table 1. Primer sequences for the confirmation of *GCNT2* deletion breakpoints in PKCC215.

Primer Pair	Direction	Sequence	Annealing Temperature (˚C)	Region	Product Size (bp)
a	F	GCCTCTGCGAAAGTGAAATG	68	10413447–10413854	407
	R	TTGCAGCTGAGAATGTTAGGC			
b	F	TCCCCTTCTATTAGCCTGCTC	68	10462475–10462795	320
	R	GGCTCTTCTGCCATGTAAGG			
c	F	GGAAATCATCCTTGCAGAGC	68	10465833–10466212	379
	R	GAGGTGCCTTTGTGTTTGTG			
d	F	TGGCTTGGGTTCTATTCTGG	68	10466314–10466622	308
	R	TCCAGAGCTGTGAGCGAATA			
e	F	TCAACCCTGACACAGTGCTT	68	10467543–10467908	365
	R	GCACATGTCACTCTGCGTTA			
f	F	GGAACCGTGAGCCAATTAAA	70	10468290–10468674	384
	R	GCAGCAGCTAACTGGTCTCC			
g	F	AGACCAGTTAGCTGCTGCAC	68	10468657–10468929	272
	R	TGCCCTGCTAATTTTTGTATTTT			
h	F	GATTACTACCAGGGCTGTGGAC	70	10468697–10468879	184
	R	GCCAGACTGGTCTTGAACTTCT			
i	F	CAGGGCTGTGGACTGTTGTA	70	10468706–10469140	434
	R	CTGAGCAGGGACACAACTA			
j	F	ACATGGCTGCCTGTCTTCA	70	10469216–10469578	362
	R	CACCATCTGTGGTAGGCAGA			
k	F	ACCACATCATCTGCTAAAACG	70	10472212–10472581	369
	R	TTGTTTTGGGTAGTGTTGACAT			
l	F	GTGCCAGCTGTGTGACTGTT	70	10488488–10488871	383
	R	CCCAGGAGACATTCAGTGCT			
m	F	TGCATTGCGTTAGCTGAATC	70	10501483–10501879	396
	R	ACACCAGGGTTCAGGTCCTA			
n	F	AGAGGTCTGGGCAATCTCTG	70	10653054–10653438	384
	R	GTTTCCCTTGGTGGCAACT			
o	F	GCACTGGGAACTTGGAAATG	70	10654655–10655048	393
	R	TGGCTCTTCCATGGACTGTT			
p	F	AAATGCTGCCTGGATGTTG	70	10656747–10657142	395
	R	TCCCAAACTGCGATCTATGA			
q	F	GTGTGCCCACATCTCTTCCT	68	10657266–10657693	427
	R	GGTTGGATTTGGACCATGAG			
r	F	CAGAGCAGGACTCCATCTCAAT	68	10657573–10657768	195
	R	GTAGAAGCAGCCAAAAAGGAAA			
s	F	AACCTGCCACCTGTTCTTGT	68	10657690–10658111	421
	R	CCCTAGCAGTGATCCCATGT			
t	F	GATCACTGCTAGGGCTGGAG	68	10658098–10658490	392
	R	GTTTCCCTCTGCACCAGAAA			
u	F	TTTCTGGTGCAGAGGGAAAC	68	10658471–10658887	416
	R	CACAGGACAAGCTCTGAAAACA			
v	F	TGTCTACTGGGGCCTGTCA	70	10659000–10659449	449
	R	CACCACGTCCGGCTACTTT			
w	F	CCACATCCACCCTGGAATTA	68	10663284–10663680	396
	R	ATGGGGTTGGAAGCCATTAT			

Note: F: forward; R: reverse.

Table 2. Clinical characteristics of individuals of PKCC215 who exhibit autosomal recessive congenital cataracts.

Individual ID	Gender	Age of onset	Age at enrollment	Visual Acuity (OD/OS)	Clinical Findings
12	F	birth	3 months	PL/PL	B/L nuclear cataracts
15	M	birth	5 years	CF/CF	B/L nuclear cataracts; B/L nystagmus
16	F	birth	8 months	CF/PL	B/L cataracts

Note: Age of onset determined by the age at which the first symptoms manifested; OD: oculus dexter; OS: oculus sinister; PL: perception of light; CF: counting fingers; B/L: bilateral.

using predesigned *Gcnt2* TaqMan expression assays (Applied Biosystems). *Gapdh* served as an endogenous internal control. The $2^{-\Delta CT}$ method was used to determine the relative expression of *Gcnt2*, normalized to *Gapdh* expression, at each developmental stage.

Results

A large consanguineous family (designated PKCC215 in our cohort), which includes six individuals who have congenital cataracts and nine unaffected individuals, was recruited from the Punjab province of Pakistan. Medical records and interviews with the family indicated that all affected individuals developed cataracts during infancy (Table 2). Ophthalmic examinations of the participating members of PKCC215, conducted with slit-lamp microscopy, detected bilateral nuclear cataracts in every affected individual.

Initially, we screened PKCC215 for linkage to all loci and genes reported to be associated with arCC with fluorescently labeled short tandem repeat (STR) markers. Homozygosity-based analysis of the resultant alleles suggested linkage to chromosome 6p markers. Maximum two-point LOD scores of 5.78 and 5.53, at the recombinant fraction θ = 0, were obtained with the markers D6S470 and D6S309, respectively (Table 3). Haplotypes of each family member were constructed with the alleles of the STR markers, which supported the results of linkage analysis, localizing the cataractous phenotype to an interval of 12.64 cM (6.22 Mb) on chromosome 6p, between D6S309 and D6S1653 (Fig 1). Every affected individual was homozygous for the D6S470 and D6S1034 markers, while unaffected individuals were either heterozygous or homozygous for the wild-type sequence (Fig 1).

The localization interval on chromosome 6p harbors *GCNT2*, a gene previously associated with cataractogenesis as well as with the adult i phenotype. The sequencing of all five exons of *GCNT2*, the exon–intron boundaries, and the 5′ and 3′ UTR regions were completed for the unaffected members of PKCC215; however, PCR reactions with the genomic DNA of all six affected individuals failed to yield specific PCR products. It is notable that the multiplex PCR reactions of the six affected individuals did generate specific products for every primer set except that of *GCNT2* (data not shown).

To further investigate the possibility of a deletion mutation at the *GCNT2* locus in PKCC215, we examined the whole-exome data of three affected individuals, which revealed

Table 3. Two-point LOD scores of PKCC215 and microsatellite markers used for linkage analysis.

| Marker	cM	Mb	0	0.01	0.05	0.09	0.1	0.2	0.3	Z_{max}	θ_{max}
D6S1574	9.18	6.01	−∞	0.89	1.86	1.95	1.94	1.50	0.91	1.95	0.09
D6S309	14.07	8.22	5.53	5.43	5.00	4.55	4.43	3.23	1.94	5.53	0.00
D6S470	18.22	10.02	5.78	5.69	5.22	4.75	4.63	3.34	2.03	5.78	0.00
D6S1034	23.23	12.2	−∞	3.41	3.64	3.41	3.33	2.38	1.34	3.64	0.05
D6S1653	26.71	14.44	−∞	-2.91	-0.90	-0.26	-0.16	0.33	0.34	0.34	0.30
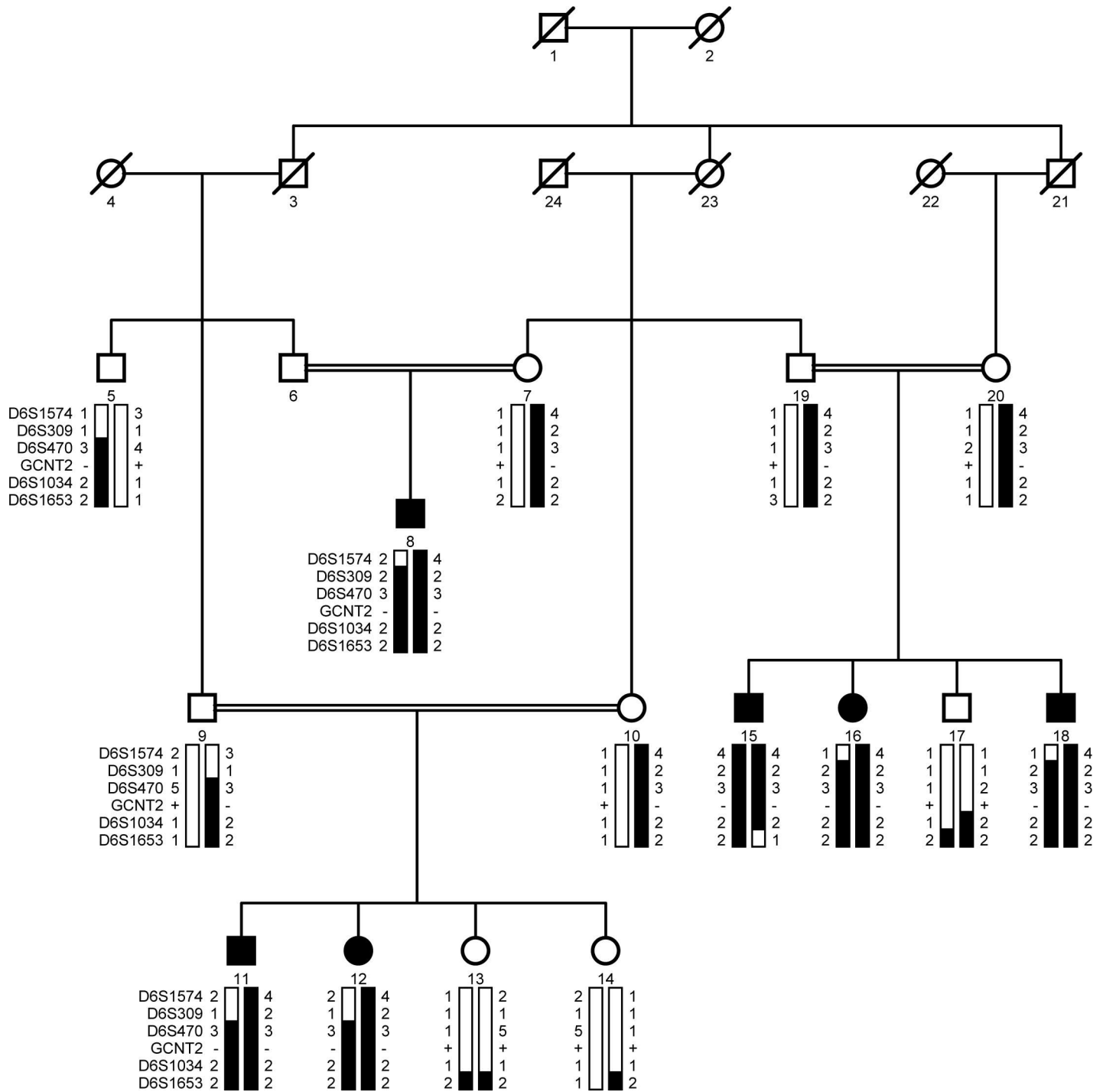
Fig 1. Pedigree diagram of the family, PKCC215. The haplotypes of 5 adjacent chromosome 6p microsatellite markers are shown. Alleles constituting the risk haplotype are shaded black, and those not segregating with cataracts are shown in white. Square: male; circle: female; filled symbol: affected; double line between symbols: consanguineous mating; diagonal line through symbol: deceased.

two absent variants on the chromosome 6p interval. The two absent SNPs, rs35318586 (chr6: 104,650,35 bp) and rs3756954 (chr6: 107,245,60 bp), identified the proximal and distal boundaries of a deletion.
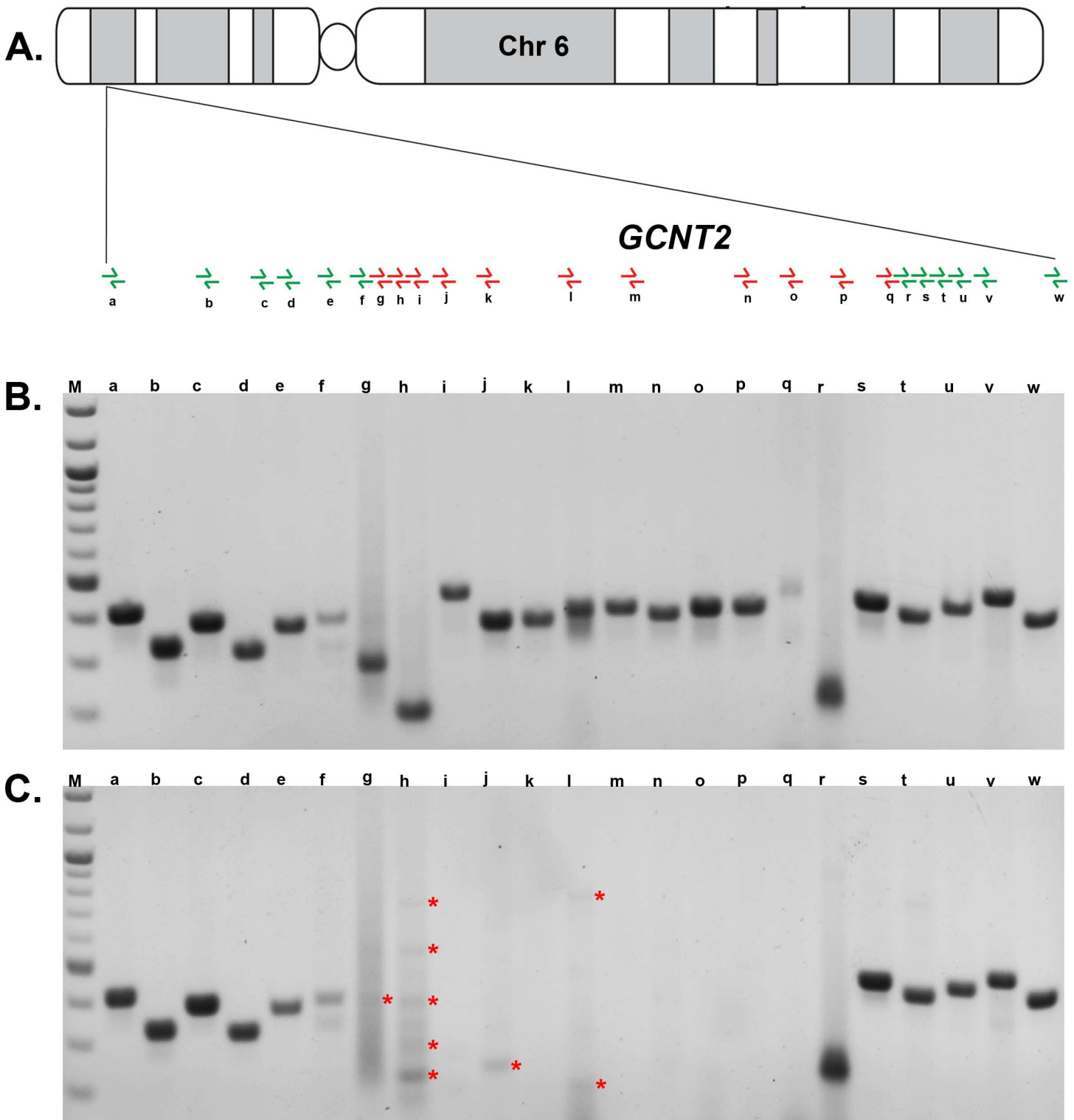
Fig 2. Characterization of the *GCNT2* deletion responsible for congenital cataracts in PKCC215. A) Model of chromosome 6 indicating the placement of the primer pairs in the PCR amplification of the *GCNT2* region. Visualization (on a 1.5% agarose gel) of the amplification products from each primer pair with the DNA of B) an unaffected and C) affected member of PKCC215. Amplification of the genomic DNA of the affected individual was successful (produced specific PCR products) with the primer pairs indicated in green. Amplification of the affected individual's DNA failed (non-specific or no PCR products) with the primer pairs indicated in red. Note: Asterisk indicates non-specific PCR products.

doi:10.1371/journal.pone.0167562.g002
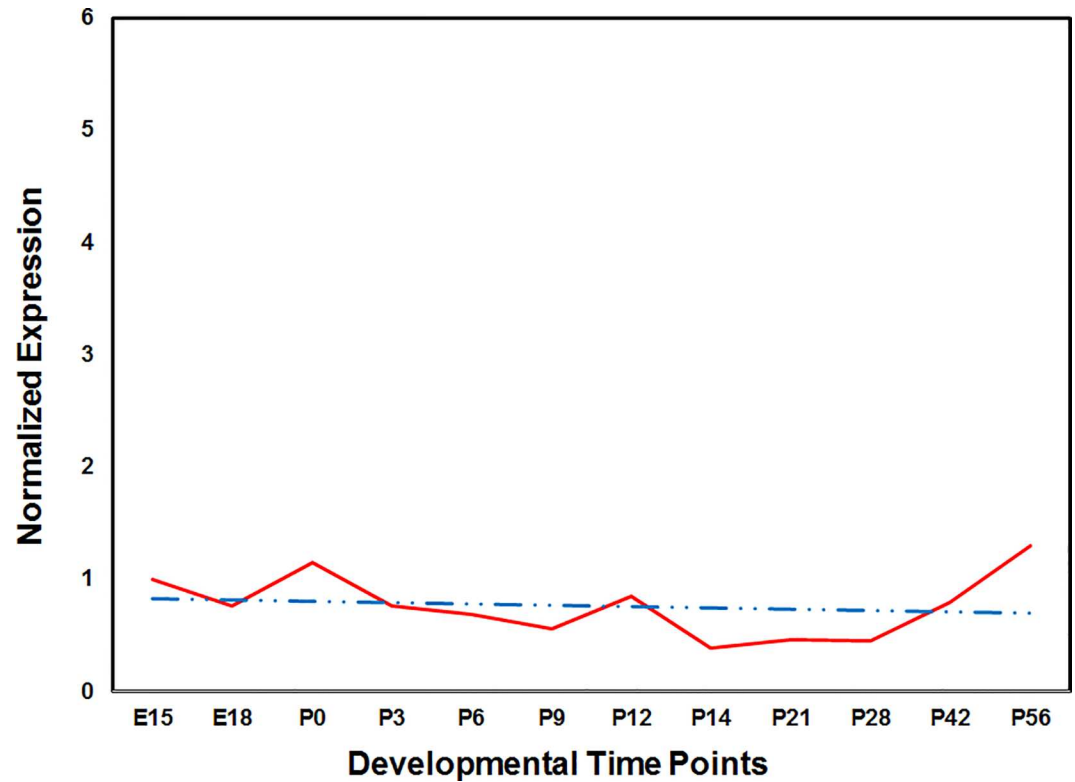
Fig 3. Expression profile of glucosaminyl (N-acetyl) transferase 2 in the developing mouse lens. The expression of *Gcnt2* at different developmental time points was normalized to *Gapdh*. The x-axis and y-axis represent the developmental time point and the normalized expression level of *Gcnt2* mRNA, respectively. Note: red line: expression of *Gcnt2*; blue dashed line: linear regression.

doi:10.1371/journal.pone.0167562.g003

To further confirm the deletion and refine the proximal and distal boundaries, we designed 23 primer pairs spanning *GCNT2* and the flanking 5′ and 3′ regions (Fig 2A and Table 1). The
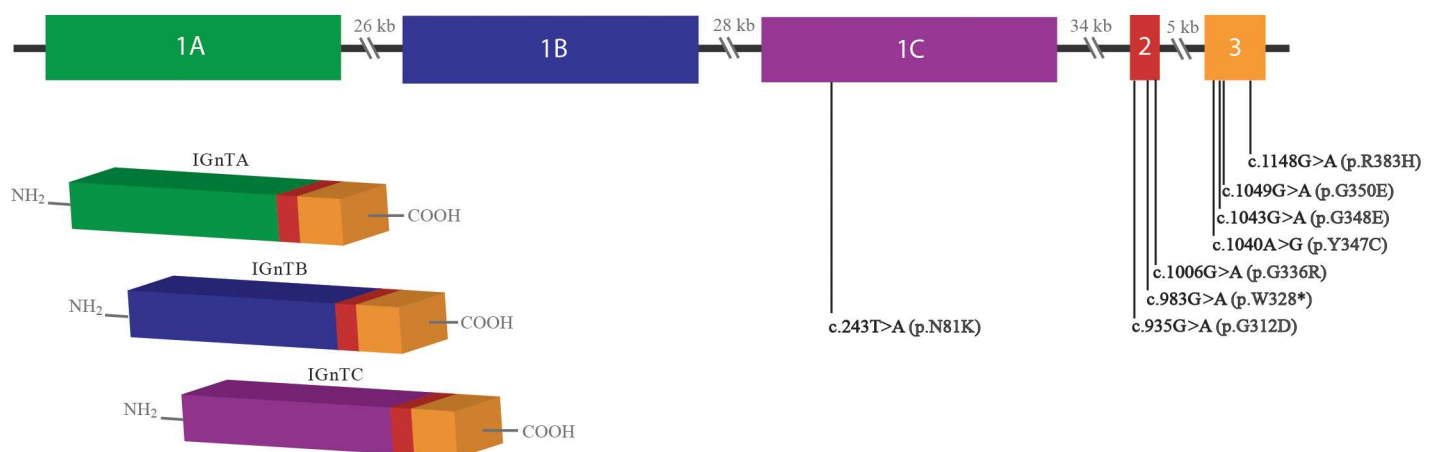
Fig 4. A schematic of the *GCNT2* isoforms generated through alternative splicing. *GCNT2* extends 135 kb on chromosome 6p24.3–24.2, and produces three isoforms through alternative splicing that differ in their first coding exon (1A, 1B, or 1C). To date, 10 different causal mutations have been identified in *GCNT2*, including seven missense mutations, one nonsense mutation, and two large genomic deletions (Table 4). Mutations residing in exon 2 or 3 generate both the adult i phenotype and congenital cataracts.

doi:10.1371/journal.pone.0167562.g004

Table 4. Mutations reported in *GCNT2*.

No.	Nucleotide Change	Amino Acid Change	Coding Exon	Phenotype	Reference
1	c.243T>A	p.N81K	1C	adult i	[50]
2	c.935G>A	p.G312D	2	arCC; adult i	[38]
3	c.983G>A	p.W328*	2	arCC; adult i	[13]
4	c.1006G>A	p.G336R	2	cataract; adult i	[37]
5	c.1040A>G	p.Y347C	3	arCC	[39]
6	c.1043G>A	p.G348E	3	arCC; adult i	[35]
7	c.1049G>A	p.G350E	3	cataract; adult i	[37]
8	c.1148G>A	p.R383H	3	arCC; adult i	[35]
9	deletion	-	1B, 1C, 2, 3	arCC; adult i	[35]
10	deletion	-	1B, 1C, 2, 3	arCC; adult i	[40]

Note: arCC: autosomal recessive congenital cataracts.

doi:10.1371/journal.pone.0167562.t004

primer pairs were used to amplify the genomic DNA of an unaffected parent, individual 7, and an affected child, individual 8. As shown in Fig 2, the PCR reactions distal to primer pair "f" and proximal to primer pair "r" failed to amplify the genomic DNA of the affected individual. Subsequently, we tested the genomic DNA of the other five affected individuals (11, 12, 15, 16, and 18) with the same 23 primer pairs. The results of the PCR reactions with the DNA of these five individuals were equivalent to those of affected individual 8; amplification reactions with primer pairs annealing to genomic DNA on chromosome 6 in the 10468657–10657693 region failed to yield specific products (data not shown). The causal deletion (10468657–10657693) on chromosome 6p harbors *GCNT2*, *RP1-290I10.5* (long intergenic non-coding RNA), *RP11-360O19.4* (antisense non-coding RNA), and *RP11-360O19.5* (a processed pseudogene). Among the deleted species, only *GCNT2* is expressed in the ocular lens, based on both the next-generation transcriptome sequencing [45,46] and proteome profiling of the developing mouse lens (unpublished data).

To identify the proximal and distal breakpoints of the deletion, we set out to amplify across the deletion with the forward primer from the "f" primer pair and the reverse primer from the "r" primer pair (the new combination was designated "LR1"). We designed an additional primer pair, "LR2", to amplify across the deletion with a forward primer annealing proximal to the forward primer of LR1 and a reverse primer annealing distal to the reverse primer of LR1. We attempted to amplify the genomic DNA of each of the six affected individuals with both the LR1 and LR2 primer pairs. None of the 12 PCR reactions, with primer pairs LR1 and LR2 or their reciprocal combinations, yielded specific products. The failure to amplify across the deletion indicates that the affected members of PKCC215 carry a complex chromosomal rearrangement at the *GCNT2* locus comprised of a large deletion accompanied, most likely, by an insertion.

As a first step in understanding the physiological significance of *GCNT2* in the development of the ocular lens, we explored the expression of *Gcnt2* in our recently published mouse lens transcriptome. *Gcnt2* is expressed as early as embryonic day 15, and low but consistent expression was present at all six time points examined in the study [45]. To further investigate the expression profile of *Gcnt2*, we performed quantitative real-time PCR analysis at six additional postnatal time points, using the expression level at E15 as a reference for the comparison of the subsequent 11 time points. As shown in Fig 3, expression levels of *Gcnt2* remain steady throughout the early developmental period. Such consistent expression suggests a critical role for *GCNT2* in the development of the lens and the maintenance of its transparency.

Discussion

Here, we report a novel deletion in *GCNT2* that causes autosomal recessive congenital cataracts. Initial exclusion analysis localized the critical interval to chromosome 6p with statistically significant two-point LOD scores. Next, PCR amplification of the coding exons of *GCNT2* suggested a large deletion at the locus. The deletion was further confirmed by a lack of chromosome 6p variants in the exomes of affected individuals. Chromosome walking delineated the deletion to approximately 190 kb on chromosome 6p, encompassing *GCNT2*, which segregates with the cataractous phenotype in the family.

GCNT2, also named *IGnT*, extends 135 kb on chromosome 6p24.3–24.2, and alternative splicing produces three isoforms that each contain a different initial coding exon (1A, 1B, or 1C), followed by exons 2 and 3, which are common to the three transcripts [35,47]. To date, 10 different causal mutations have been identified in *GCNT2*, including seven missense mutations, one nonsense mutation, and two large genomic deletions (Fig 4 and Table 4). Interestingly, mutations residing in the first exon result in the adult i phenotype without congenital cataracts whereas the mutations in exons 2 and 3 result in both the adult i phenotype and congenital cataracts. Yu and colleagues recently proposed a partial association of the two traits, the adult i phenotype, and congenital cataracts. They suggested that a defect in only the C isoform, IGnTC, leads to the reduction of I antigen in red blood cells, whereas a deficiency in the function of all three IGnT enzymes results in cataractogenesis [48].

In 2001, Yu and colleagues published three Taiwanese pedigrees with the adult i phenotype; each member of the three families who exhibited the adult i phenotype also manifested congenital cataracts [35]. The expression of the I antigen in blood is regulated only by the IGnTC isoform of *GCNT2*. Therefore, mutations in exon 1A or 1B have not been associated with the adult i phenotype [35,47]. In their report, Yu and colleagues suggested that the association between the adult i phenotype and congenital cataracts may be due to a functional dependency of I-related and cataract-related genes, or might be a pleiotropic effect of the adult i phenotype gene on cataractogenesis [48].

Chen and colleagues developed a mouse model lacking the third exon of *Gcnt2* to investigate the physiological function of the enzyme [49]. However, the *Gcnt2* knockout mice did not develop cataracts during early development and, although older mice manifested mild cataracts in the cortical regions, there was no difference in the degree of opacity between the lenses of wild-type and knockout mice [49]. Nevertheless, the genetic data associating cataractogenesis with mutations in *GCNT2* remains strong. The discrepancy between the cataractous phenotypes between patients with *GCNT2* mutations and that of the *Gcnt2*-deficient mice might be due to tolerance or compensation for *GCNT2*-encoded glycosyltransferases. Li and colleagues suggested that differences in the function or expression of other glycosyltransferases in mice might be responsible for the lack of cataractogenesis in the *Gcnt2* knockout [48].

In conclusion, we have identified a large deletion at the *GCNT2* locus spanning ~190 kb that is responsible for autosomal recessive congenital cataracts in a large family. This is the largest deletion reported at the *GCNT2* locus and is likely accompanied by an insertion. Identifying the proteins, such as GCNT2, that are essential for the maintenance of lens transparency is vital for understanding and preventing cataractogenesis. The identification of new genes associated with cataractogenesis will help us to better understand the biology of the ocular lens at a molecular level.

Supporting Information

S1 File. The ARRIVE checklist has been completed for the manuscript.
(PDF)

Acknowledgments

Author Contributions

References

1. Wride MA (2011) Lens fibre cell differentiation and organelle loss: many paths lead to clarity. Philos Trans R Soc Lond B Biol Sci 366: 1219–1233. 366/1568/1219 [pii]. doi: 10.1098/rstb.2010.0324 PMID: 21402582

2. Bermejo E, Martinez-Frias ML (1998) Congenital eye malformations: clinical-epidemiological analysis of 1,124,654 consecutive births in Spain. Am J Med Genet 75: 497–504. 10.1002/(SICI)1096-8628 (19980217)75:5<497::AID-AJMG8>3.0.CO;2-K [pii]. PMID: 9489793

3. Apple DJ, Ram J, Foster A, Peng Q (2000) Elimination of cataract blindness: a global perspective entering the new millenium. Surv Ophthalmol 45 Suppl 1: S1–196.

4. Gilbert C, Foster A (2001) Childhood blindness in the context of VISION 2020—the right to sight. Bull World Health Organ 79: 227–232. PMID: 11285667

5. Shiels A, Bennett TM, Hejtmancik JF (2010) Cat-Map: putting cataract on the map. Mol Vis 16: 2007–2015. PMID: 21042563

6. Shiels A, Hejtmancik JF (2013) Genetics of human cataract. Clin Genet 84: 120–127. doi: 10.1111/cge.12182 PMID: 23647473

7. Gill D, Klose R, Munier FL, McFadden M, Priston M, Billingsley G et al. (2000) Genetic heterogeneity of the Coppock-like cataract: a mutation in CRYBB2 on chromosome 22q11.2. Invest Ophthalmol Vis Sci 41: 159–165. PMID: 10634616

8. Butt T, Yao W, Kaul H, Xiaodong J, Gradstein L, Zhang Y et al. (2007) Localization of autosomal recessive congenital cataracts in consanguineous Pakistani families to a new locus on chromosome 1p. Mol Vis 13: 1635–1640. PMID: 17893665

9. Kaul H, Riazuddin SA, Shahid M, Kousar S, Butt NH, Zafar AU et al. (2010) Autosomal recessive congenital cataract linked to EPHA2 in a consanguineous Pakistani family. Mol Vis 16: 511–517. PMID: 20361013

10. Ponnam SP, Ramesha K, Tejwani S, Ramamurthy B, Kannabiran C (2007) Mutation of the gap junction protein alpha 8 (GJA8) gene causes autosomal recessive cataract. J Med Genet 44: e85. doi: 10.1136/jmg.2007.050138 PMID: 17601931

11. Pras E, Pras E, Bakhan T, Levy-Nissenbaum E, Lahat H, Assia EI et al. (2001) A gene causing autosomal recessive cataract maps to the short arm of chromosome 3. Isr Med Assoc J 3: 559–562. PMID: 11519376

12. Sabir N, Riazuddin SA, Butt T, Iqbal F, Nasir IA, Zafar AU et al. (2010) Mapping of a new locus associated with autosomal recessive congenital cataract to chromosome 3q. Mol Vis 16: 2634–2638. PMID: 21179239

13. Pras E, Raz J, Yahalom V, Frydman M, Garzozi HJ, Pras E et al. (2004) A nonsense mutation in the glucosaminyl (N-acetyl) transferase 2 gene (GCNT2): association with autosomal recessive congenital cataracts. Invest Ophthalmol Vis Sci 45: 1940–1945. PMID: 15161861

14. Kaul H, Riazuddin SA, Yasmeen A, Mohsin S, Khan M, Nasir IA et al. (2010) A new locus for autosomal recessive congenital cataract identified in a Pakistani family. Mol Vis 16: 240–245. PMID: 20161816

15. Sabir N, Riazuddin SA, Kaul H, Iqbal F, Nasir IA, Zafar AU et al. (2010) Mapping of a novel locus associated with autosomal recessive congenital cataract to chromosome 8p. Mol Vis 16: 2911–2915. PMID: 21203409

16. Heon E, Paterson AD, Fraser M, Billingsley G, Priston M, Balmer A et al. (2001) A progressive autosomal recessive cataract locus maps to chromosome 9q13-q22. Am J Hum Genet 68: 772–777. doi: 10.1086/318798 PMID: 11179024

17. Safieh LA, Khan AO, Alkuraya FS (2009) Identification of a novel CRYAB mutation associated with autosomal recessive juvenile cataract in a Saudi family. Mol Vis 15: 980–984. 103 [pii]. PMID: 19461931

18. Smaoui N, Beltaief O, BenHamed S, M'Rad R, Maazoul F, Ouertani A et al. (2004) A homozygous splice mutation in the HSF4 gene is associated with an autosomal recessive congenital cataract. Invest Ophthalmol Vis Sci 45: 2716–2721. doi: 10.1167/iovs.03-1370 PMID: 15277496

19. Yasmeen A, Riazuddin SA, Kaul H, Mohsin S, Khan M, Qazi ZA et al. (2010) Autosomal recessive congenital cataract in consanguineous Pakistani families is associated with mutations in GALK1. Mol Vis 16: 682–688. PMID: 20405025

20. Riazuddin SA, Yasmeen A, Zhang Q, Yao W, Sabar MF, Ahmed Z et al. (2005) A new locus for autosomal recessive nuclear cataract mapped to chromosome 19q13 in a Pakistani family. Invest Ophthalmol Vis Sci 46: 623–626. doi: 10.1167/iovs.04-0955 PMID: 15671291

21. Pras E, Levy-Nissenbaum E, Bakhan T, Lahat H, Assia E, Geffen-Carmi N et al. (2002) A missense mutation in the LIM2 gene is associated with autosomal recessive presenile cataract in an inbred Iraqi Jewish family. Am J Hum Genet 70: 1363–1367. doi: 10.1086/340318 PMID: 11917274

22. Ramachandran RD, Perumalsamy V, Hejtmancik JF (2007) Autosomal recessive juvenile onset cataract associated with mutation in BFSP1. Hum Genet 121: 475–482. doi: 10.1007/s00439-006-0319-6 PMID: 17225135

23. Pras E, Frydman M, Levy-Nissenbaum E, Bakhan T, Raz J, Assia EI et al. (2000) A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. Invest Ophthalmol Vis Sci 41: 3511–3515. PMID: 11006246

24. Riazuddin SA, Yasmeen A, Yao W, Sergeev YV, Zhang Q, Zulfiqar F et al. (2005) Mutations in betaB3-crystallin associated with autosomal recessive cataract in two Pakistani families. Invest Ophthalmol Vis Sci 46: 2100–2106. doi: 10.1167/iovs.04-1481 PMID: 15914629

25. Chen J, Ma Z, Jiao X, Fariss R, Kantorow WL, Kantorow M et al. (2011) Mutations in FYCO1 cause autosomal-recessive congenital cataracts. Am J Hum Genet 88: 827–838. S0002-9297(11)00201-1 [pii]. doi: 10.1016/j.ajhg.2011.05.008 PMID: 21636066

26. Aldahmesh MA, Khan AO, Mohamed JY, Alghamdi MH, Alkuraya FS (2012) Identification of a truncation mutation of acylglycerol kinase (AGK) gene in a novel autosomal recessive cataract locus. Hum Mutat 33: 960–962. doi: 10.1002/humu.22071 PMID: 22415731

27. Lachke SA, Alkuraya FS, Kneeland SC, Ohn T, Aboukhalil A, Howell GR et al. (2011) Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. Science 331: 1571–1576. 331/6024/1571 [pii]. doi: 10.1126/science.1195970 PMID: 21436445

28. Zhao L, Chen XJ, Zhu J, Xi YB, Yang X, Hu LD et al. (2015) Lanosterol reverses protein aggregation in cataracts. Nature 523: 607–611. nature14650 [pii]. doi: 10.1038/nature14650 PMID: 26200341

29. Cohen D, Bar-Yosef U, Levy J, Gradstein L, Belfair N, Ofir R et al. (2007) Homozygous CRYBB1 deletion mutation underlies autosomal recessive congenital cataract. Invest Ophthalmol Vis Sci 48: 2208–2213. doi: 10.1167/iovs.06-1019 PMID: 17460281

30. Fukuda M, Fukuda MN, Hakomori S (1979) Developmental change and genetic defect in the carbohydrate structure of band 3 glycoprotein of human erythrocyte membrane. J Biol Chem 254: 3700–3703. PMID: 438154

31. Watanabe K, Hakomori SI, Childs RA, Feizi T (1979) Characterization of a blood group I-active ganglioside. Structural requirements for I and i specificities. J Biol Chem 254: 3221–3228. PMID: 107170

32. Wiener AS, Unger LJ, Cohen L, Feldman J (1956) Type-specific cold auto-antibodies as a cause of acquired hemolytic anemia and hemolytic transfusion reactions: biologic test with bovine red cells. Ann Intern Med 44: 221–240. PMID: 13292836

33. Marsh WL, Jenkins WJ (1960) Anti-i: a new cold antibody. Nature 188: 753.

34. Fukuda M, Carlsson SR, Klock JC, Dell A (1986) Structures of O-linked oligosaccharides isolated from normal granulocytes, chronic myelogenous leukemia cells, and acute myelogenous leukemia cells. J Biol Chem 261: 12796–12806. PMID: 3462190

35. Yu LC, Twu YC, Chang CY, Lin M (2001) Molecular basis of the adult i phenotype and the gene responsible for the expression of the human blood group I antigen. Blood 98: 3840–3845. PMID: 11739194

36. Yamaguchi H, Okubo Y, Tanaka M (1972) A note on possible close linkage between the Ii blood locus and a congenital cataract locus. Proc. Jpn. Acad. 48: 625–628, 1972.

37. Inaba N, Hiruma T, Togayachi A, Iwasaki H, Wang XH, Furukawa Y et al. (2003) A novel I-branching beta-1,6-N-acetylglucosaminyltransferase involved in human blood group I antigen expression. Blood 101: 2870–2876. 2002-09-2838 [pii]. doi: 10.1182/blood-2002-09-2838 PMID: 12468428

38. Wussuki-Lior O, Abu-Horowitz A, Netzer I, Almer Z, Morad Y, Goldich Y et al. (2011) Hematologic biomarkers in childhood cataracts. Mol Vis 17: 1011–1015. 113 [pii]. PMID: 21541272

39. Aldahmesh MA, Khan AO, Mohamed JY, Hijazi H, Al-Owain M, Alswaid A et al. (2012) Genomic analysis of pediatric cataract in Saudi Arabia reveals novel candidate disease genesGenet Med 14: 955–962. gim201286 [pii]. doi: 10.1038/gim.2012.86 PMID: 22935719

40. Borck G, Kakar N, Hoch J, Friedrich K, Freudenberg J, Nurnberg G et al. (2012) An Alu repeat-mediated genomic GCNT2 deletion underlies congenital cataracts and adult i blood group. Hum Genet 131: 209–216. doi: 10.1007/s00439-011-1062-1 PMID: 21761136

41. Jiaox X, Khan SY, Irum B, Khan AO, Wang Q, Kabir F et al. (2015) Missense Mutations in CRYAB Are Liable for Recessive Congenital Cataracts. PLoS One 10: e0137973. PONE-D-15-12822 [pii]. doi: 10.1371/journal.pone.0137973 PMID: 26402864

42. Jiao X, Kabir F, Irum B, Khan AO, Wang Q, Li D et al. (2016) A Common Ancestral Mutation in CRYBB3 Identified in Multiple Consanguineous Families with Congenital Cataracts. PLoS One 11: e0157005. PONE-D-15-23789 [pii]. doi: 10.1371/journal.pone.0157005 PMID: 27326458

43. Lathrop GM, Lalouel JM (1984) Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36: 460–465. PMID: 6585139

44. Schaffer AA, Gupta SK, Shriram K, Cottingham RW (1994) Avoiding recomputation in genetic linkage analysis. Hum Hered 44: 225–237. PMID: 8056435

45. Khan SY, Hackett SF, Lee MC, Pourmand N, Talbot CC Jr., Riazuddin SA (2015) Transcriptome Profiling of Developing Murine Lens Through RNA Sequencing. Invest Ophthalmol Vis Sci 56: 4919–4926. 2422128 [pii]. doi: 10.1167/iovs.14-16253 PMID: 26225632

46. Khan SY, Hackett SF, Riazuddin SA (2016) Non-coding RNA Profiling of Developing Murine Lens. Exp Eye Res. S0014-4835(16)30009-4 [pii].

47. Yu LC, Twu YC, Chou ML, Reid ME, Gray AR, Moulds JM et al. (2003) The molecular genetics of the human I locus and molecular background explain the partial association of the adult i phenotype with congenital cataracts. Blood 101: 2081–2088. 2002-09-2693 [pii]. doi: 10.1182/blood-2002-09-2693 PMID: 12424189

48. Yu LC, Lin M (2011) Molecular genetics of the blood group I system and the regulation of I antigen expression during erythropoiesis and granulopoiesis. Curr Opin Hematol 18: 421–426. doi: 10.1097/MOH.0b013e32834baae9 PMID: 21912254

49. Chen GY, Muramatsu H, Kondo M, Kurosawa N, Miyake Y, Takeda N et al. (2005) Abnormalities caused by carbohydrate alterations in Ibeta6-N-acetylglucosaminyltransferase-deficient mice. Mol Cell Biol 25: 7828–7838. 25/17/7828 [pii]. doi: 10.1128/MCB.25.17.7828-7838.2005 PMID: 16107727

50. Lin M, Hou MJ, Yu LC (2006) A novel IGnT allele responsible for the adult i phenotype. Transfusion 46: 1982–1987. TRF01006 [pii]. doi: 10.1111/j.1537-2995.2006.01006.x PMID: 17076854