BMC Ophthalmology

RESEARCH Open Access

Identification of mutations associated with congenital cataracts in nineteen Chinese families

Hai-Sen Sun[1,2†], Teng Huang[1,2†], Zhe-Xuan Liu[1,2], Yi-Tong Xu[1,2], Ya-Qi Wang[1,2], Ming-Cheng Wang[4], Shen-Rong Zhang[1,2], Jia-Lin Xu[1,2], Kai-Yi Zhu[1,2], Wen-Kai Huang[1,2], Xiu-Feng Huang[3*] and Jin Li[1,2*]

Abstract

Background Congenital cataracts (CC) are one of the leading causes of impaired vision or blindness in children, with approximately 8.3–25% being inherited. The aim of this study is to investigate the mutation spectrum and frequency of 9 cataract-associated genes in 19 Chinese families with congenital cataracts.

Purpose To identify the gene variants associated with congenital cataracts.

Methods This study included a total of 58 patients from 19 pedigrees with congenital cataracts. All probands were initially screened by whole-exome sequencing(WES), and then validated by co-segregation analysis using Sanger sequencing.

Results Likely pathogenic variants were detected in 8 families, with a positivity rate of 42.1%. Variants in various genes were identified, including *GJA3*, *CRYGD*, *CRYBA4*, *BFSP2*, *IARS2*, *CRYAA*, *CRYBA1*, *ARL2* and *CRYBB3*. Importantly, this study identified compound heterozygous variants of *IARS2* in one family.

Conclusions Our research findings have revealed multiple gene variants associated with cataracts, providing clinical guidance for improved molecular diagnosis of congenital cataracts in the era of precision medicine.

Keywords Congenital cataract, Whole-exome sequencing, Variants

Background

Congenital cataracts (CC) are one of the leading causes of visual impairment or blindness in children [1]. Approximately 8.3–25% of congenital cataracts are hereditary [2], and there are around 314,000 new cases of pediatric cataracts (congenital and developmental) each year in the world. If left untreated, congenital cataracts might result in irreversible visual impairment [3]. At present, sequence variants in more than 50 genes are associated with inherited forms of isolated or primary cataracts-largely based on the current update of the Cat-Map database (https://cat-map.wustl.edu/, accessed on 12 June 2024) [4]. Variants associated with cataracts typically occur in genes

†Hai-Sen Sun and Teng Huang contributed equally to this work.

Jin Li and Xiu-Feng Huang contributed equally as co-corresponding authors.

*Correspondence:
Xiu-Feng Huang
hxfwzmc@163.com
Jin Li
lijin_wz@eye.ac.cn
[1]State Key Laboratory of Ophthalmology, Optometry and Visual Science, Eye Hospital, Wenzhou Medical University, Wenzhou, China
[2]National Clinical Research Center for Ocular Diseases, Eye Hospital, Wenzhou Medical University, Wenzhou, China
[3]Zhejiang Provincial Clinical Research Center for Pediatric Disease, The Second Affiliated Hospital, Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China
[4]Institute of PSI Genomics Co., Ltd, Wenzhou, China

that encode various crystallin proteins—including crystallins (*CRYAA*, *CRYAB*, *CRYBA1*, *CRYBA2*, *CRYBA4*, *CRYBB1*, *CRYBB2*, *CRYBB3*, *CRYGA*, *CRYGB*, *CRYGC*, *CRYGD*, *and CRYGS*), gap junction proteins (*GJA3* and *GJA8*), intermediate filament proteins, membrane proteins (*MIP* and *LIM2*), cytoskeleton proteins (*BFSP2* and *VIM*), developmental regulators, and transcription factors (*MAF*, *PITX3*, *HSF4*, *PAX6*, *FOXE3*, and *EYA1*) [2, 5–7]. Because of the extreme genetic heterogeneity of CC, molecular diagnosis is challenging for ophthalmologists and researchers.

Whole-exome sequencing (WES) is a powerful approach for detecting exome variants. In the past decade, WES has become an important method for identifying pathogenic genes in monogenic diseases [8, 9]. According to the National Disabled Persons' Database of the China Disabled Persons' Federation, there were approximately 1,109,000 visually impaired children aged 0–14 in China in 2021, with around 69,000 in Zhejiang Province. In this study, we aimed to perform WES to identify disease-causing variants in probands with CC from the southern region of Zhejiang Province, China.

This study recruited 19 families, comprising a total of 58 cases of CC, at the Eye and Vision Hospital affiliated with Wenzhou Medical University from April 2021 to April 2023. Through WES and comprehensive bioinformatics analyses, we identified variants in various genes, including *GJA3*, *CRYGD*, *BFSP2*, *IARS2*, *CRYAA*, *CRYBA1*, and *ARL2*. Additionally, two variants classified as "Variants of Uncertain Significance (VUS)" were identified in *CRYBA4* and *CRYBB3*. Importantly, this study identified a novel variant in *IARS2*. Our findings expanded the mutation spectrum associated with congenital cataracts, facilitating early diagnosis and treatment. Targeted Whole-exome sequencing in inherited congenital cataract patients provided significant diagnostic information.

Methods

Subjects and sample collection

This study was conducted with the approval of the Ethics Committee of Wenzhou Medical University Eye Hospital (approval number 2021-239-k-209) and adhered to the guidelines of the Helsinki Declaration for researching familial relationships and DNA specimens. Each participant underwent a detailed written informed consent procedure. To ensure the accuracy and reliability of the study, all patients and their family members received comprehensive ocular examinations by professional ophthalmologists, including visual acuity, slit-lamp biomicroscopy, ophthalmoscopy, corneal evaluations, axial length measurements (IOL Master-700; Carl Zeiss Meditec AG), and fundus examination for both eyes (Table 1 and Supp. Table S1). With the cooperation of the patients and their family members, peripheral blood samples or buccal mucosal samples were collected for further research analysis. Genomic DNA was extracted using a QIAGEN Blood DNA extraction kit (QIAGEN, Germany) or the Thermo Fisher Invitrogen™ kit (MagMAX™ DNA Multi-Sample Ultra 2.0 Kit, Thermo Fisher Scientific, Norway), according to the manufacturer's instructions.

Whole-exome sequencing and bioinformatics analysis

To conduct the analysis, DNA samples obtained from individuals affected by the condition underwent comprehensive sequencing of the entire exome (WES). This process utilized the Twist Human Core Exome Kit (Twist Bioscience, USA) and the NovaSeq 6000 platform (Illumina, San Diego, USA) for the sequencing procedure. The short-read sequence data were aligned to the hg19 human reference genome using the Burrows–Wheeler Aligner tool (BWA) [10] and variant calling with GATK [11] according to best practice guidelines. Variant annotation was performed using ANNOVAR [12], In silico prediction tools Combined Annotation-Dependent Depletion (CADD v1.7) [13], REVEL [14], SpliceAI [15], SIFT [16], MutationTaster2021 [17], PolyPhen-2 [18] and AlphaMissence [19] used to predict potential pathogenic variations; Since CCs are rare in the Chinese population, only variants with a frequency below 0.5% were chosen. The population allele frequencies of identified variants were cross-verified using data from Genome Aggregation Database (gnomAD v4.1.0) data [20], China Metabolic Analytics Project (ChinaMAP) data [21] and Westlake BioBank for Chinese (WBBC) data [22]. Then disease and phenotype databases including Online Mendelian Inheritance in Man (OMIM; http://www.omim.org), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar), the Human Gene Mutation Database (HGMD; http://www.hgmd.org), and Human Phenotype Ontology (HPO; https://hpo.jax.org/app/), were used for variant interpretation. The above databases were accessed in June 2023.

Variants previously described as disease-causing in the Human Gene Mutation Database (HGMD) and literature were given the highest priority. The effects of mutations described in the HGMD were further validated by reviewing published literature reporting the variants. Only those variants with convincing evidence were selected. In addition, protein-truncation mutations, such as nonsense and frameshift (insertions or deletions), were also ranked higher in priority. Variants were analyzed for possible pathogenic significance according to the 2015 American College of Medical Genetics and Genomics (ACMG) guidelines [23–26].

Sanger sequencing validation

Sanger sequencing validation was performed by synthesizing primers targeting the DNA fragment of interest,

Table 1 Clinical features of affected probands with variants identified in this study

Family (Proband)	Age of onset	Cataract description	Procedures/age	IOP	HCD/age	AXL/age	BCVA (logMAR)	Additional clinical fingdings
F#1	5y	Lamellar cataract	BL Lensectomy + IOL/5y	R18.9 L21.8(PTS) R17.3 L14.1(PO 6 m)	R11.8 L11.7/5y	R21.47 L21.53/5y	R0.5 L0.7(PST) R0.1 L0.1(PO 1y)	Nil
F#2	2 m	Total cataract	BL Lensectomy/4m BL IOL/3y	R16.67 L16.77(PTS) R18.3 L17.1(PO 2y)	N/K	R21.26 L21.40/5y	R0.3 L0.6(PO 5y)	BL Nystagmus Dxotropia
F#3	35y	Embryonic nuclear cataract	L Lensectomy + IOL/41y	R13.5 L13.5(PTS) R17.4 L14.4(PO 1d)	R11.5 L11.4/41y	R22.99 L22.99/41y	R0.2 L FC/BE(PST) R0.2 L1.3(PO 1d)	BL Shallow Anterior Chamber BL Ciliary Body Cyst
F#4	5y	Posterior polar cataract	BL Lensectomy + IOL/5y	R16.4 L18.4(PTS) R12.1 L12.2(PO 1 m)	R11.4 L11.5/8y	R20.63 L21.28/8y	R0.8 L0.7(PST) R0.1 L0.1(PO 3y)	BL Amblyopia
F#5	5y	Lamellar cataract	BL Lensectomy + IOL/5y	R16.8 L18.0(PTS) R14.0 L16.0(PO 1y)	R11.7 L11.3/5y	R21.79 L21.84/8y	R0.8 L0.7(PST) R0.2 L0.1(PO 1y)	BL Amblyopia
F#6	1y	Embryonic nuclear cataract	BL Lensectomy + IOL/3y	N/K	R10.2 L10.3/3y	R19.63 L19.72/6y	R0.4 L0.2(PO 1y)	BL Visual impairment Exotropia
F#7	6y	Embryonic nuclear cataract	Nil	R16.6 L17.1	R11.7 L11.8/16y	R21.96 L21.97/16y	R0.1 L0.1	Nil
F#8	3y	Embryonic nuclear cataract	BL Lensectomy + IOL/4y	R13.9 L14.2(PO 2y)	R10.4 L10.4/6y	R22.62 L22.29/6y	R0.9 L0.7(PST) R0.2 L0.1(PO 1y)	R Exotropia
F#9	8y	anterior polar cataract	BL Lensectomy + IOL/39y	R12.3 L11.3(PTS) R14.0 L11.3(PO 1d)	R11.6 L11.6/39y	R22.64 L22.86/39y	R0.7 L0.8(PST)	BL Corneal endothelial dystrophy
F#10	16y	anterior polar cataract	Nil	R11.4 L10.7	N/K	N/K	R0 L0	BL Shallow Anterior Chamber

F Family, L Left, R Right, SB since birth, BL bilateral, PC posterior chamber, IOL Intraocular Lens, IOP intraocular pressure, HCD Horizontal Corneal Diameter in mm and age of measurement, AXL axial length in mm and age of measurement, BCVA best corrected visual acuity, FC Finger Count, BE before Eye, N/K not known, nil empty, y year, m month, d day. PTS prior to surgery, PO postoperation

followed by PCR amplification. Sanger sequencing was then carried out using an ABI 3730xl sequencer (Applied Biosystems, USA). The obtained sequences were compared with the reference sequence using MutationMapper software.

Results

Clinical findings

All patients in this study had various types of congenital cataracts and no other systemic diseases. Other ophthalmic findings of the nineteen probands were listed in Table 1 and Supp.Table S1. Phenotypes of the probands with congenital cataracts were recorded (Table 1). However, two probands (Family 2: III 2, Family 9: II5) had undergone cataract surgery prior to this study, and the remaining two probands (Family 7: III 1, Family 10: II5) had not undergone surgery and did not have anterior segment photographs available. The phenotypes of these families were determined based on their medical records.

Identification of suspected causative variants

This study included a total of 58 patients from 19 pedigrees with congenital cataracts. Among these patients, likely pathogenic variants were detected in 8 pedigrees, with a positivity rate of 42.1%. Of the 8 variants in this study, the variant of *IARS2* (c.1067G > T: p.G356V) was novel, with the remaining 7 variants already reported (Table 2). Additionally, two variants classified as "VUS" were identified in Family 3 and Family 10. Most families exhibited autosomal dominant inheritance, with only Family 5 being considered as having autosomal recessive inheritance. Notably, likely causative variants were identified in *GJA3* in two families, and one family each for *CRYGD*, *BFSP2*, *IARS2*, *CRYAA*, *CRYBA1* and *ARL2* (Table 2). The specific results of the family co-segregation verification within the pedigrees in this study were detailed in Figs. 1, 2 and 3.

The prediction results of the variants by SIFT, Mutation Taster2021, PolyPhen-2, CADD, REVEL, SpliceAI and AlphaMissence were listed (Table 2). The pathogenic variants of family 1, 2, 4, 5, 6, 7, 8, and 9 are shown by sequencing chromatograph (Figs. 1 and 2), and

Table 2 Analysis of likely pathogenic variants in familial and sporadic cataracts

Family	Gene	Transcript/ position	DNA Change	Predicted amino acid change	Mutation Type	Inheritance	ACMG Classification	SIFT/Mut-Taster/ PolyPhen	CADD_ Phred	REVEL	SpliceAI	AlphaMissence	GnomAD4_ exome	ChinaMAP/ WBBC	HGMD	Novel
1	*GJA3*	NM_021954.4 (chr13:20143113))	c.176 C > T	p.P59L	nonsynonymous SNV	AD	LP (PS4 + PM1 + PM2 + PP1 + PP3)	D/D/D	29.4	0.97	0.0	0.86	0	0/0	DM	Bennett, et al. 2004
2	*CRYGD*	NM_006891.4 (chr2208121780)	c.418 C > T	p.R140*	stopgain	AD	P (PVS1 + PS3 + PS4 + PM2 + PM6 + PP1 + PP4)	-	36	-	0.0	-	0	0/0	DM	Jackson, et al. 2020
3	*CRYBA4*	NM_001886.3 (chr22:26630360)	c.464 C > T	p.P155L	nonsynonymous SNV	N/K	VUS (PM1 + PM2 + PP3)	D/D/D	26.5	0.79	0.0	0.68	5.5e-6	0/0	-	Yes
4	*BFSP2*	NM_003571.4 (chr3:13348608–13348610)	c.697_699del	p.E233del	non-frameshift deletion	AD	LP (PS4 + PM2 + PM4 + PP1)	-	-	-	0.0	-	6.8e-7	0/0	DM	Zhang, et al. 2004
5	*IARS2*	NM_018060.4 (chr1:220142958)	c.1067G > T	p.G356V	nonsynonymous SNV	AR	P (PM1 + PM2 + PM3)	- /D/D	33	0.69	0.13	0.75	0	0/0	-	Yes
	IARS2	NM_018060.4 (chr1:220105891)	c.2575T > C	p.F859L	nonsynonymous SNV		VUS (PM1 + PM2 + PP3)	- /D/D	31	0.38	0.0	0.97	1.9e-5	7.8e-4/7.8e-4	DM	Li, et al. 2018
6	*CRYAA*	NM_000394.4 (chr21:44589270)	c.61 C > T	p.R21W	nonsynonymous SNV	AD	P (PS3 + PS4 + PM2 + PM5 + PP1 + PP3)	D/D/D	29.6	0.89	0.0	0.86	0	0/0	DM	Hansen, et al.2007
7	*GJA3*	NM_021954.4 (chr13:20142700)	c.589 C > T	p.P197S	nonsynonymous SNV	AD	LP (PM1 + PM2 + PP1 + PP3 + PP4)	D/D/D	27.8	0.97	0.01	0.96	0	0/0	DM	Ponnam et al. 2013
8	*CRYBA1*	NM_005208.5 (chr17:29242117–29242119)	c.272_274del	p.G91del	non-frameshift deletion	AD	P (PS2 + PS3 + PS4 + PM1 + PM2 + PM4 + PP1)	-	-	-	0.0	-	1.4e-6	0/0	DM	Lu, et al. 2007
9	*ARL2*	NM_001667.4 (chr11:65014251)	c.44G > T	p.R15L	nonsynonymous SNV	AD	LP (PS3 + PM2 + PM6 + PP1)	D/D/B	27.3	0.99	0.0	0.68	3.05-5	1.7e-3/1.7e-3	DM	Cai, et al. 2019
10	*CRYBB3*	NM_004076.5 (chr22:25201254)	c.-20-123G > C	-	splicing variant	N/K	VUS (PM2 + PM3)	-	12	-	0.23	-	7.2e-6	1.5e-4/1.1e-4	-	Yes

Abbreviations: N/K not known; D, damaging; T, tolerated; B, benign; HGMD, Human Gene Mutation Database; DM, Disease-causing Mutation; ACMG, American College of Medical Genetics and Genomics; P, pathogenic; LP, likely pathogenic; VUS, Variant of uncertain significance; "-", Data Not Available; GnomAD4_exome: GnomAD All individuals exome allele frequency; ChinaMAP: China Metabolic Analytics Project allele frequency; WBBC: Westlake BioBank for Chinese allele frequency

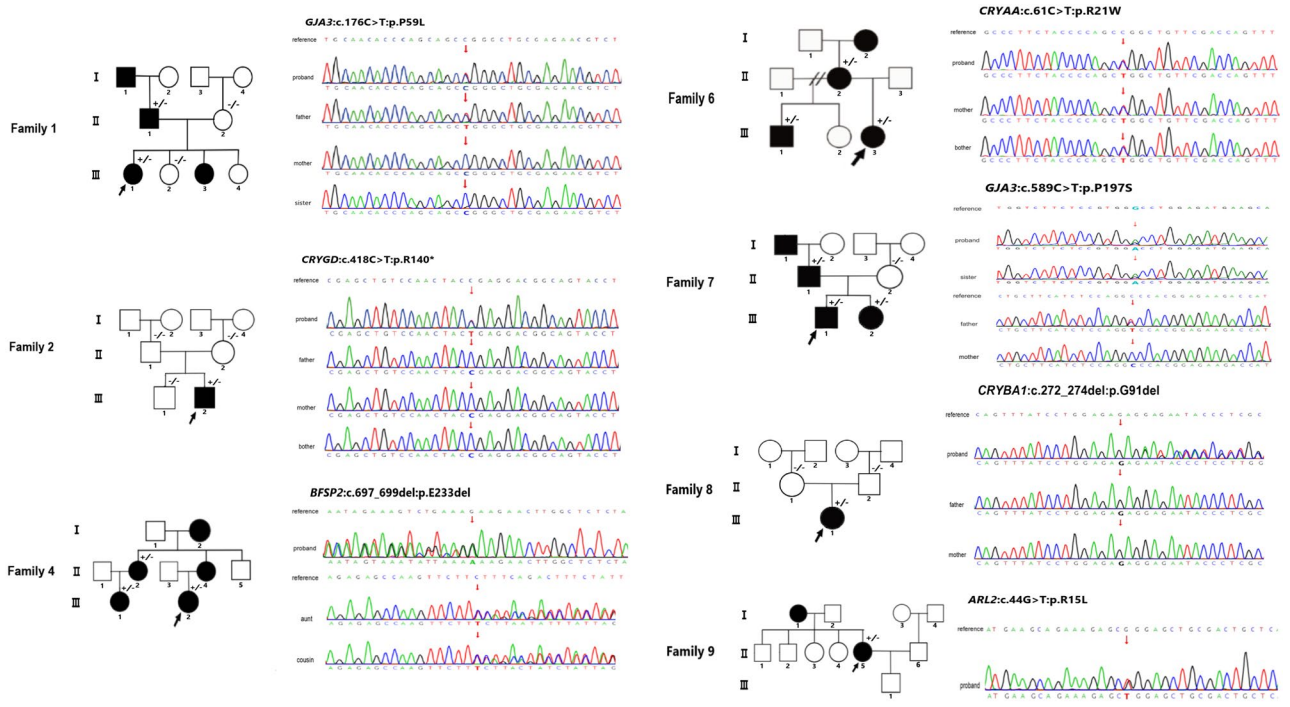

Fig. 1 Pedigrees and chromatograms of families 1, 2, 4, 6, 7, 8, and 9 with mutations. Squares indicate men and circles women; black and white symbols represent affected and unaffected individuals, respectively. The proband is marked with an arrow, and +/– indicates heterozygous individuals, –/– indicates individuals testing negative

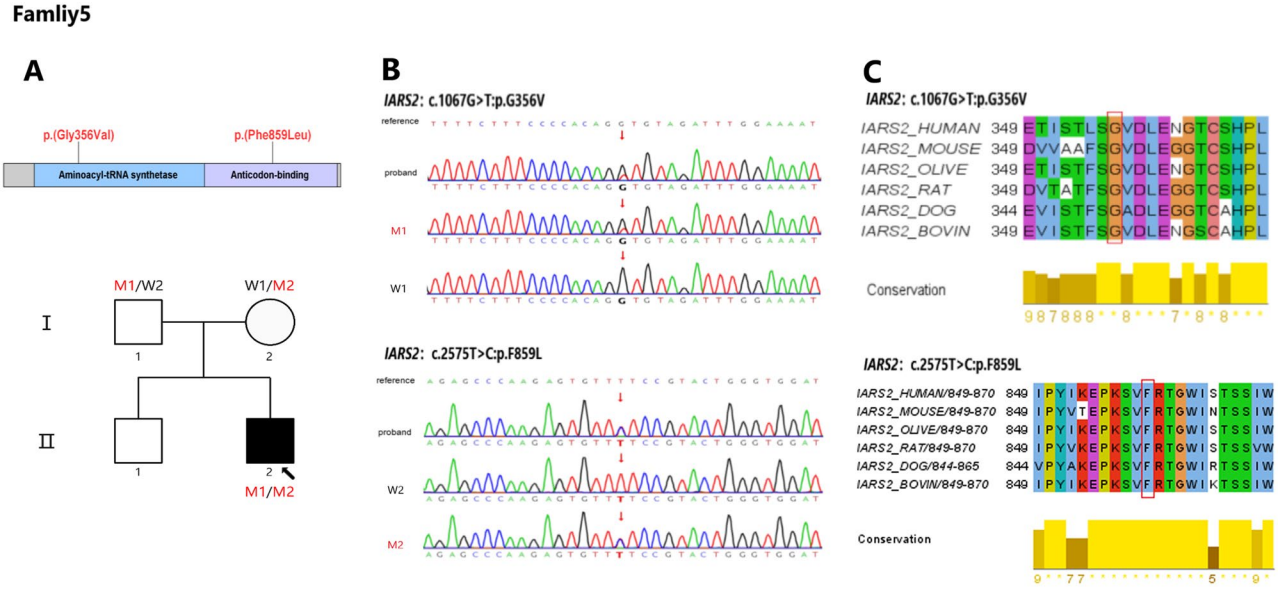

Fig. 2 Pedigree and variants in *IARS2*. The schematic show the encoded domain structure of IARS2 (**A-B**). The variants found in this study are illustrated above the schematics. W1: Wild Type 1, W2: Wild Type 2, M1: Mutant Type 1, M2: Mutant Type 2; The multiple-sequence alignments from different species. (**C**): Arrows show conserved glycine at p.G356 and Phenylalanine at p.F859 in *IARS2* protein. Abbreviation: OLIVE, Olive Baboon;

the sequencing chromatograph of family 3 and 10 are depicted using Integrative Genomics Viewer (IGV) plots (Fig. 3). Given that the *IARS2* (c.1067G > T: p.G356V) is novel, we performed multiple protein sequence alignments using T-Coffee (https://www.ebi.ac.uk/jdispatcher/msa/tcoffee/) [27] and Jalview (Version: 2.11.4.0) [28] to demonstrate that the variant is highly conserved across different species (Fig. 2C) [28]. To further support the pathogenicity of these variants, we constructed a comprehensive table of population mutation frequencies,

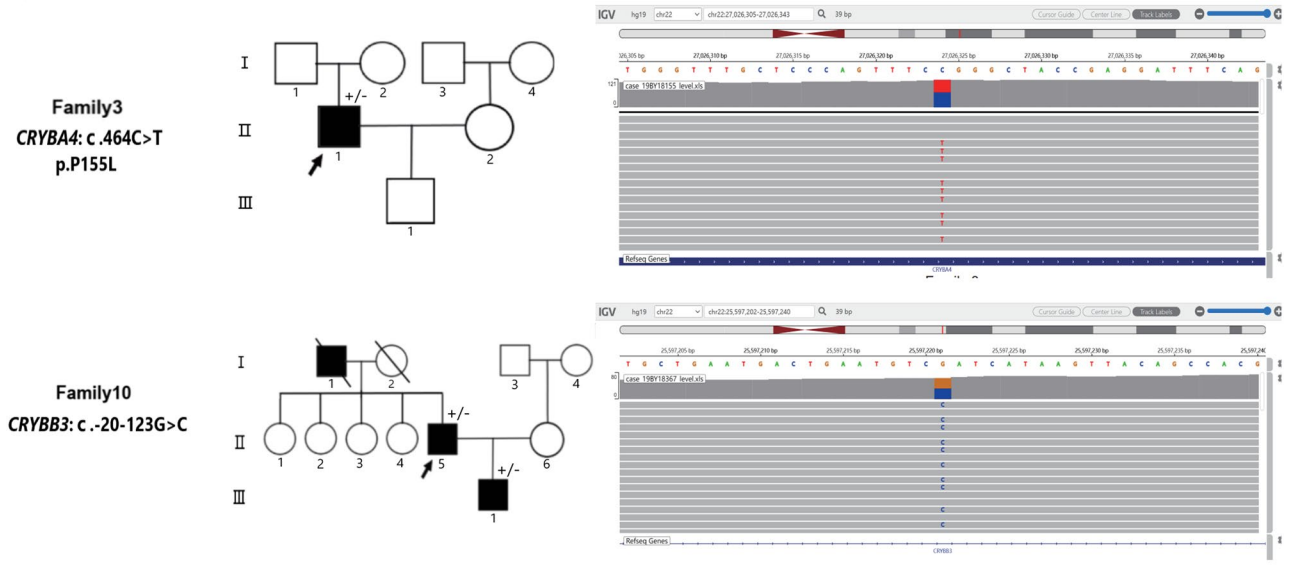

Fig. 3 Pedigrees and IGV plots of families 3 and 10 with mutations. Squares indicate men and circles women; black and white symbols represent affected and unaffected individuals, respectively. The proband is marked with an arrow, and +/− indicates heterozygous individuals

including GnomAD, ChinaMAP, and WBBC (Table 2). No causative variants were identified in the remaining nine pedigrees among the cataract-related genes examined. The corresponding pedigrees of these nine families with unidentified variants are presented in Supp.Fig S1.

Five of nineteen mutations were identified in crystallin genes, while the remaining five mutations were identified in four genes. Among these, two mutations were in *GJA3* (MIM: 121001), one was in cytoskeletal protein (*BFSP2*; MIM:107320), and the other two were located in the *IARS2* (MIM: 612801) and *ARL2* (MIM: 610170), respectively. In the ten mutations, six are missense mutations, one is a nonsense mutation, two are non-frameshift mutations, and one is an intronic mutation. In this study, the compound heterozygous mutation in *IARS2* (c.1067G > T; p.G356V) identified in Family 5 was considered a novel variant (Table 2).

Discussion

Here, we performed targeted exome sequencing on probands from 19 families with congenital cataracts. It was discovered that 10 cases contained putative pathogenic variants in nine cataract-related genes, including missense mutations, nonsense mutations, non-frameshift deletions, and splicing mutation (Table 2). Additionally, we have elaborated on the types of cataracts and other ocular clinical manifestations involved in these families (Table 1). Therefore, by performing genetic testing on these patients from diverse families, our aim is to discover genetic variations related to the development of congenital cataracts. This will contribute to a better understanding of hereditary congenital cataracts and offer clinical guidance for better molecular diagnosis of congenital cataracts in the era of precision medicine.

Two variants were detected in the *GJA3* gene, exon2: c.176 C > T in Family 1 and exon2: c.589 C > T in family 7, both of which have been previously reported [5, 29]. The variant c.176 C > T in family 1 was classified as likely pathogenic (LP) and was detected in multiple pedigrees with CCs. Li Wang et al. [30] reported that this variant (c.176 C > T) co-segregated with CCs in three affected family members. The c.589 C > T variant in Family 7 was also classified as likely pathogenic (LP). The variant (c.589 C > T) co-segregated with CCs in six affected family members [29]. The human *GJA3* gene encodes a protein of 435 amino acids, located on chromosome 13q12.11 [31]. The Connexin 46 (Cx46) protein, encoded by *GJA3*, is primarily expressed in lens fiber cells and plays a critical role in maintaining lens transparency, particularly in mature fiber cells at the central core of the lens [32].

In this study, we have identified mutations associated with the crystallin proteins *CRYGD*, *CRYAA*, and *CRYBA1* in family 2, family 6, and family 8, respectively. To date, a total of 308 disease-causing variants in crystallins have been identified, accounting for nearly 23.0% of all inherited cataract variants [2, 33]. The variant *CRYGD*:exon3:c.418 C > T in family 2 is classified as a pathogenic (P) variant and has been shown to co-segregate with disease in seven affected family members from two pedigrees [34, 35]. Cellular experiments have suggested that this variant causes impaires gene function [36]. The *CRYGD* gene that is mutated in this family encodes the crystallin protein γD, which is an important component of the lens. It possesses a specific protein

structure that contributes to maintaining the normal function of the lens and plays a crucial role in lens transparency and normal visual function [37]. In family 6, the c.61 C > T:p.R21W variant was detected in the *CRYAA* gene of the proband's brother and mother. This variant has been classified as P according to the American College of Medical Genetics (ACMG) guidelines. Several published studies have previously detected this variant in patients from families affected by cataracts [38–42], and functional experiments have indicated that this variant affects protein function [43–46]. The *CRYAA* gene is an important lens structural protein primarily present in lens fiber cells. It plays a role in maintaining lens transparency and stabilizing lens structure. Gene variants in *CRYAA* can potentially lead to endoplasmic reticulum stress (ERS)-induced unfolded protein response (UPR) during lens development, causing cellular apoptosis and subsequently leading to cataract formation [47]. The variant *CRYBA1*:exon4:c.272_274del in family 8 is classified as P, and has been shown to segregate with disease in 16 affected family members from three pedigrees [48, 49]. Cell transfection experiments further suggest that this variant leads to impaired gene function [49]. Variants in the *CRYBA1* gene may lead to abnormalities in protein structure, preventing proper folding and assembly and thereby impacting the normal development and transparency of the lens. In addition to hereditary cataracts, *CRYBA1* variants may also be associated with other ocular abnormalities such as corneal lesions and retinal diseases [50].

The *BFSP2* gene encodes a protein called beaded filament structural protein 2 (*BFSP2*), which serves as an important structural support in the lens of the eye and other tissues and plays a vital role in maintaining the normal structure and function of the eye and nervous system [51]. The variant *BFSP2*: exon3: c.697_699del in family 4 was classified as LP and has been reported to segregate with disease in multiple affected family members [52, 53]. This variant is a deletion mutation that results in a change in the length of the protein without causing a change to the open reading frame [54]. The cataract phenotype resulting from the *BFSP2*: c.697_699del is mostly nuclear, Y-sutural, stellate, or spokelike cortical, which is consistent with the posterior polar radial cataract reported in family 4 (Fig. 4) [53]. The specific pathogenesis remains to be further fundamental research.

In family 5, a novel compound heterozygous mutation was identified in *IARS2*. Specifically, the exon9: c.1067G > T was classified as VUS, and the exon21: c.2575T > C was classified as LP. This gene is a nuclear gene that encodes mitochondrial isoleucyl-tRNA synthetase [55]. The variants *IARS2* (c.1067G > T: p.G356V and c.2575T > C: p.F859L) of the proband (II2) were inherited from the father and mother respectively, with both parents being heterozygous for the variants (Fig. 2A-B). Variants in the *IARS2* gene typically result in an autosomal recessive inherited disease known as CAGSSS (cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss, and skeletal dysplasia; OMIM: 616007). Two variants were detected in *IARS2*, which encodes an aminoacyl-tRNA synthetase subunit, an important enzyme involved in protein synthesis. The presence of these variants may impact the normal function of the *IARS2* gene, thereby adversely affecting the process of protein synthesis. Furthermore, studies have suggested that variants in the *IARS2* gene are associated with various genetic disorders, including cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss, and skeletal developmental abnormalities [55]. The p.G356V lies in the class Ia aminoacyl-tRNA synthetases domain, while the p.F859L localizes to the anticodon-binding domain (Fig. 2A). Multiple sequence alignments indicated that Glycine (G) at position 356 and Phenylalanine (F) at position 859 of IARS2 were highly conserved among different species (Fig. 2C). Although this variant has not been reported in previous studies, based on the location of the variant and the nature of the base change, it can be inferred that this variant might lead to structural or functional changes in the protein, thereby causing the development of diseases such as cataracts.

The variant *ARL2*: exon1:c.44G > T in family 9 was classified as LP. Since we were unable to collect the genetic data of the parents, co-segregation verification could not be performed for this family. Nevertheless, Xue-Bi Cai et al. [56] reported this variant in a Chinese pedigree, where it co-segregated with four affected family members. Additionally, they demonstrated that this variant impairs gene function through a transgenic mouse model. In their report, the proband and her three daughters exhibited microcornea, retinal cone dystrophy, cataracts, and posterior scleral staphyloma, which are consistent with the features of MRCS syndrome. However, in Family 9, we did not observe any of these symptoms except for cataracts. The *ARL2* (ADP-ribosylation factor-like 2) gene encodes a small GTPase that plays an important role in intracellular signaling, endocytosis, and regulation of the cellular cytoskeleton (*ARL2*_HUMAN, P36404). The variant identified in this gene may lead to changes in protein structure or function, thereby affecting the normal signal transduction and metabolic processes within cells and potentially causing the development of cataracts.

According to the guidelines of the ACMG, the significance of the CRYBA4 variant in family 3 and the CRYBB3 variant in family 10 was classified as "Variants of Uncertain Significance". The variant CRYBA4:exon6:c.464 C > T: p.P155L found in family 3 was a novel variant at this specific locus that is classified as of uncertain significance.

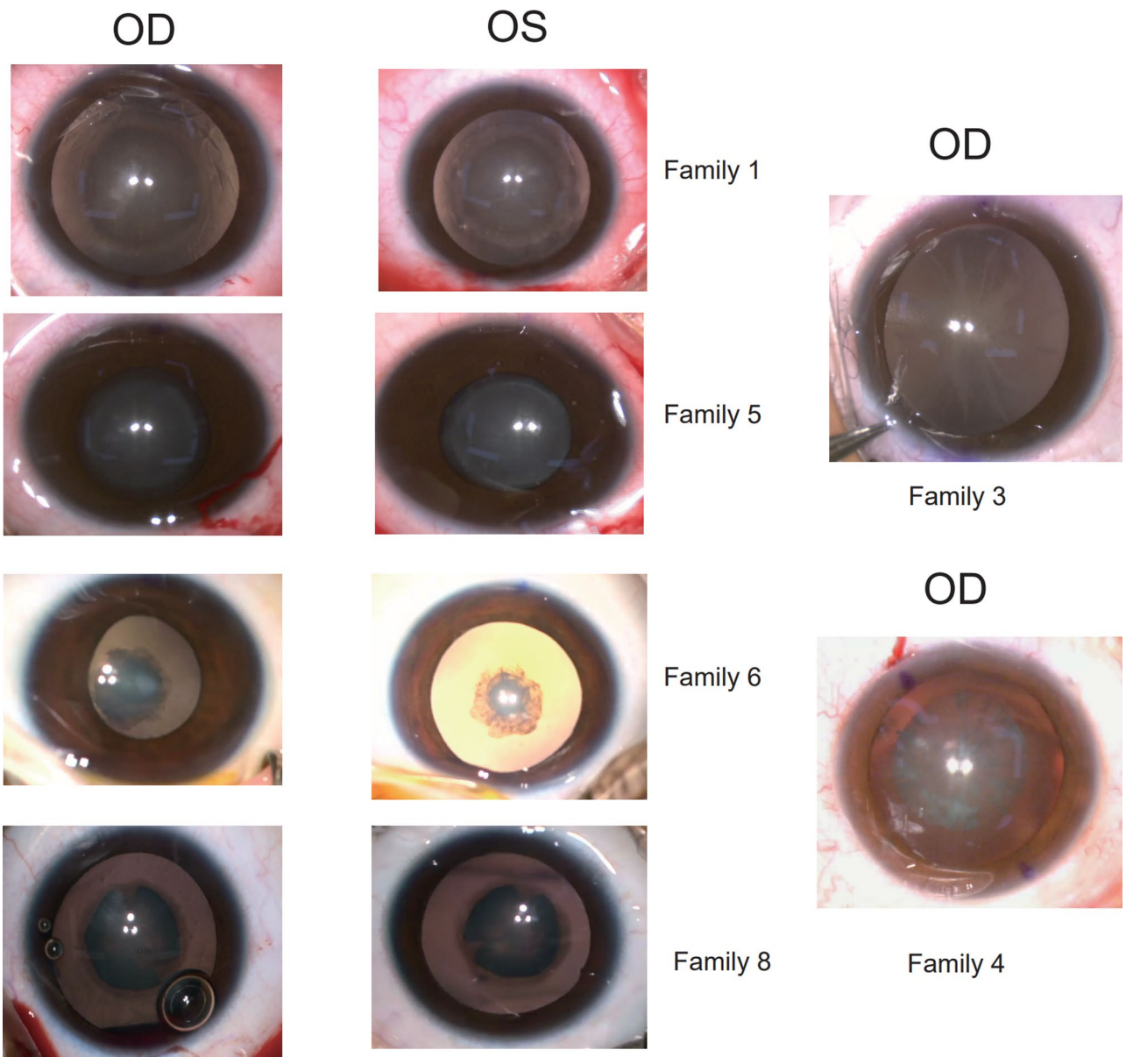

Fig. 4 Phenotype of nuclear cataract in families 1, 3, 4, 5, 6, and 8. Abbreviations: OD OCULUSDEXTER (Right eye), OS OCULUSSINISTER (Left eye)

This variant occurs in the beta/gamma-crystallin functional domain (PM1). This variant was not detected in ChinaMAP and WBBC. The frequencies of this variant in gnomAD were 5.5e-06 (PM2). According to various statistical methods (such as SIFT, MutationTaster, and REVEL), this variant is predicted to have deleterious effects on the gene or gene product (PP3). In family 10, the variant CRYBB3: intron1: c.-20-123G>C was a novel variant identified at this specific locus, and it is classified as of uncertain significance. The Population allele frequencies of this variant in GnomAD, ChniaMAP and WBBC were 7.2e-6, 1.5e-4 and 1.1e-4, respectively (PM2). The prediction results from SpliceAI showed that this variant may affect splicing (PP3) (Table 2). Our limitation lies in the fact that we failed to acquire the genetic collection data of the proband's parents in family 3 for the verification of family co-segregation. In family 10, Whole Exome Sequencing (WES) indicates that the variant in CRYBB3 was detected in both the proband and the father, and both individuals are affected by congenital anterior polar cataracts (Fig. 2). However, the evidence for pathogenicity based on the current information of mutation sites in family 3 and family 10 was still insufficient, and further functional experiments are required to verify whether these mutation sites are pathogenic.

Conclusion

In conclusion, our findings revealed multiple gene variants associated with cataracts in 10 of the 19 families studied, providing clinical guidance for improved molecular diagnosis of congenital cataracts in the era of precision medicine. Additionally, large-scale exome sequencing and genetic variation analysis can enhance a deeper understanding of the pathogenesis and genetic patterns of congenital cataracts, providing a scientific foundation for the development of new therapeutic approaches and preventive measures.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12886-025-03920-4.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
We thank Catherine Perfect, MA (Cantab), from Liwen Bianji (Edanz) (www.liwenbianji.cn), for editing the English text of a draft of this manuscript.

Author contributions
HSS, XFH, and JL were involved in designing the study.TH, YTX, YQW, and SRZ contributed to collecting samples and clinical data. HSS, TH, JLX, KYZ, WKH, XFH, and JL participated in interpreting the data. HSS, TH, HXF, and JL wrote the main text of the manuscript. ZXL, MCW contributed to the revision of the article. All authors reviewed and approved the final version of the manuscript.

Funding
This study was supported by the National Natural Science Foundation of China (Grant No. 81670840).

Data availability
All data generated or analyzed during this study are included in this published article. And sequencing data have been submitted to NCBI SRA database: PRJNA1035489.

Declarations

Ethical approval
This study has been approved by the human ethics committee of wenzhou medical university laboratory (approval number 2021-239-k-209) and adheres to the provisions of the helsinki declaration for the protection of human subjects. written informed consent has been obtained from all participants involved in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 March 2024 / Accepted: 13 February 2025
Published online: 25 February 2025

References

1. Zhuang X, Wang L, Song Z, Xiao W. PLoS ONE. 2015;10(7):e0131471.
2. Shiels A, Hejtmancik JF. Mutations and mechanisms in congenital and age-related cataracts. Exp Eye Res. 2017;156:95–102.
3. Sheeladevi S, Lawrenson JG, Fielder AR, Suttle CM. Global prevalence of childhood cataract: a systematic review. Eye (Lond). 2016;30(9):1160–9.
4. Shiels A. Through the cat-map gateway: a brief history of Cataract Genetics. Genes (Basel). 2024;15(6):785.
5. Bennett TM, Mackay DS, Knopf HL, Shiels A. A novel missense mutation in the gene for gap-junction protein alpha3 (GJA3) associated with autosomal dominant nuclear punctate cataracts linked to chromosome 13q. Mol Vis. 2004;10:376–82.
6. Ma AS, Grigg JR, Ho G, et al. Sporadic and familial congenital cataracts: Mutational Spectrum and New diagnoses using next-generation sequencing. Hum Mutat. 2016;37(4):371–84.
7. Li D, Wang S, Ye H, et al. Distribution of gene mutations in sporadic congenital cataract in a Han Chinese population. Mol Vis. 2016;22:589–98.
8. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for mendelian disease gene discovery. Nat Rev Genet. 2011;12(11):745–55.
9. Aldahmesh MA, Khan AO, Mohamed JY, et al. Genomic analysis of pediatric cataract in Saudi Arabia reveals novel candidate disease genes. Genet Med. 2012;14(12):955–62.
10. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
11. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491–8.
12. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
13. Schubach M, Maass T, Nazaretyan L, Röner S, Kircher M. CADD v1.7: using protein language models, regulatory CNNs and other nucleotide-level scores to improve genome-wide variant predictions. Nucleic Acids Res. 2024;52(D1):D1143–54.
14. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an Ensemble Method for Predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877–85.
15. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from primary sequence with deep learning. Cell. 2019;176(3):535–e54824.
16. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat Protoc. 2016;11(1):1–9.
17. Steinhaus R, Proft S, Schuelke M, Cooper DN, Schwarz JM, Seelow D. MutationTaster2021. Nucleic Acids Res. 2021;49(W1):W446–51.
18. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248–9.
19. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science. 2023;381(6664):eadg7492.
20. Gudmundsson S, Singer-Berk M, Watts NA, et al. Variant interpretation using population databases: lessons from gnomAD. Hum Mutat. 2022;43(8):1012–30.
21. Cao Y, Li L, Xu M, et al. The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals. Cell Res. 2020;30(9):717–31.
22. Cong PK, Bai WY, Li JC, et al. Genomic analyses of 10,376 individuals in the Westlake BioBank for Chinese (WBBC) pilot project. Nat Commun. 2022;13(1):2939.
23. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–24.
24. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018;39(11):1517–24.
25. Biesecker LG, Harrison SM, ClinGen Sequence Variant Interpretation Working Group. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. Genet Med. 2018;20(12):1687–8.
26. Ghosh R, Harrison SM, Rehm HL, et al. Updated recommendation for the benign stand-alone ACMG/AMP criterion. Hum Mutat. 2018;39(11):1525–30.
27. Madeira F, Madhusoodanan N, Lee J, et al. Using EMBL-EBI services via web interface and programmatically via web services. Curr Protoc. 2024;4(6):e1065.
28. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189–91.

29. Ponnam SP, Ramesha K, Matalia J, Tejwani S, Ramamurthy B, Kannabiran C. Mutational screening of Indian families with hereditary congenital cataract. Mol Vis. 2013;19:1141–8.
30. Wang L, Chen Y, Chen X, Sun X. Further evidence for P59L mutation in GJA3 associated with autosomal dominant congenital cataract. Indian J Ophthalmol. 2016;64(7):508–12.
31. Willecke K, Jungbluth S, Dahl E, Hennemann H, Heynkes R, Grzeschik KH. Six genes of the human connexin gene family coding for gap junctional proteins are assigned to four different human chromosomes. Eur J Cell Biol. 1990;53(2):275–80.
32. Kumar NM, Gilula NB. The gap junction communication channel. Cell. 1996;84(3):381–8.
33. Berry V, Georgiou M, Fujinami K, Quinlan R, Moore A, Michaelides M. Inherited cataracts: molecular genetics, clinical features, disease mechanisms and novel therapeutic approaches. Br J Ophthalmol. 2020;104(10):1331–7.
34. Devi RR, Yao W, Vijayalakshmi P, Sergeev YV, Sundaresan P, Hejtmancik JF. Crystallin gene mutations in Indian families with inherited pediatric cataract. Mol Vis. 2008;14:1157–70.
35. Granzier JJ, Valsecchi M. Variations in daylight as a contextual cue for estimating season, time of day, and weather conditions. J Vis. 2014;14(1):22. [pii].
36. Vendra VP, Agarwal G, Chandani S, Talla V, Srinivasan N, Balasubramanian D. Structural integrity of the Greek key motif in βγ-crystallins is vital for central eye lens transparency. PLoS ONE. 2013;8(8):e70336.
37. Héon E, Priston M, Schorderet DF, et al. The gamma-crystallins and human cataracts: a puzzle made clearer. Am J Hum Genet. 1999;65(5):1261–7.
38. Kondo Y, Saitsu H, Miyamoto T, et al. Pathogenic mutations in two families with congenital cataract identified with whole-exome sequencing. Mol Vis. 2013;19:384–9.
39. Sun Z, Zhou Q, Li H, Yang L, Wu S, Sui R. Mutations in crystallin genes result in congenital cataract associated with other ocular abnormalities. Mol Vis. 2017;23:977–86.
40. Javadiyan S, Craig JE, Souzeau E, et al. Recurrent mutation in the crystallin alpha A gene associated with inherited paediatric cataract. BMC Res Notes. 2016;9:83.
41. Zhang XH, Da Wang J, Jia HY, et al. Mutation profiles of congenital cataract genes in 21 northern Chinese families. Mol Vis. 2018;24:471–7.
42. Laurie KJ, Dave A, Straga T, et al. Identification of a novel oligomerization disrupting mutation in CRYAA associated with congenital cataract in a South Australian family. Hum Mutat. 2013;34(3):435–8.
43. Raju I, Oonthonpan L, Abraham EC. Mutations in human αA-crystallin/sHSP affect subunit exchange interaction with αB-crystallin. PLoS ONE. 2012;7(2):e31421.
44. Kore R, Hedges RA, Oonthonpan L, Santhoshkumar P, Sharma KK, Abraham EC. Quaternary structural parameters of the congenital cataract causing mutants of αA-crystallin. Mol Cell Biochem. 2012;362(1–2):93–102.
45. Raju I, Abraham EC. Congenital cataract causing mutants of αA-crystallin/sHSP form aggregates and aggresomes degraded through ubiquitin-proteasome pathway. PLoS ONE. 2011;6(11):e28085.
46. Panda AK, Nandi SK, Chakraborty A, Nagaraj RH, Biswas A. Differential role of arginine mutations on the structure and functions of α-crystallin. Biochim Biophys Acta. 2016;1860(1 Pt B):199–210.
47. Jia ZK, Fu CX, Wang AL, Yao K, Chen XJ. Cataract-causing allele in CRYAA (Y118D) proceeds through endoplasmic reticulum stress in mouse model. Zool Res. 2021;42(3):300–9.
48. Lu S, Zhao C, Jiao H, et al. Two Chinese families with pulverulent congenital cataracts and deltaG91 CRYBA1 mutations. Mol Vis. 2007;13:1154–60.
49. Li D, Jing Q, Jiang Y. The identification and characterization of the p.G91 deletion in CRYBA1 in a Chinese family with congenital cataracts. BMC Med Genet. 2019;20(1):153.
50. 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 βB1/βA3-crystallin heteromer but not the βB1-crystallin homomer. Hum Mutat. 2011;32(3):E2050–60.
51. Sandilands A, Wang X, Hutcheson AM, et al. Bfsp2 mutation found in mouse 129 strains causes the loss of CP49′and induces vimentin-dependent changes in the lens fibre cell cytoskeleton. Exp Eye Res. 2004;78(4):875–89.
52. Zhang Q, Guo X, Xiao X, Yi J, Jia X, Hejtmancik JF. Clinical description and genome wide linkage study of Y-sutural cataract and myopia in a Chinese family. Mol Vis. 2004;10:890–900.
53. Jakobs PM, Hess JF, FitzGerald PG, Kramer P, Weleber RG, Litt M. Autosomal-dominant congenital cataract associated with a deletion mutation in the human beaded filament protein gene BFSP2. Am J Hum Genet. 2000;66(4):1432–6.
54. Li J, Leng Y, Han S, et al. Clinical and genetic characteristics of Chinese patients with familial or sporadic pediatric cataract. Orphanet J Rare Dis. 2018;13(1):94.
55. Schwartzentruber J, Buhas D, Majewski J, et al. Mutation in the nuclear-encoded mitochondrial isoleucyl-tRNA synthetase IARS2 in patients with cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, and peripheral neuropathy or with Leigh syndrome. Hum Mutat. 2014;35(11):1285–9.
56. Cai XB, Wu KC, Zhang X, et al. Whole-exome sequencing identified ARL2 as a novel candidate gene for MRCS (microcornea, rod-cone dystrophy, cataract, and posterior staphyloma) syndrome. Clin Genet. 2019;96(1):61–71.

Publisher's note