Supplementary data

Table S1 Genome Sequencing QC data

Sample	Total read count	% reads mapped	Average target read coverage	Average sequencing depth ‡	Average read length
CRCH4-01	976.6 M	49.3%	22	18.3 X	148 bp
CRCH5-01 †	933.8 M	98.4%	43	38.0 X	148 bp
CRCH12-01	964.5 M	79.1%	36	29.6 X	148 bp
CRCH14-04	996.1 M	79.4%	37	32.2 X	148 bp
CRCH21-05	518.2 M	81.6%	33	18.3 X	236 bp
CRCH27-01 †	981.5 M	81.5%	37	32.9 X	148 bp
CRCH28-01	965.9 M	95.4%	43	38.8 X	148 bp
CRCH29-03	2072.8 M	50.4%	49	42.7 X	147 bp
CTAS34-04	937.5 M	99.1%	43	38.4 X	148 bp
CRCH38-01	900.9 M	88.4%	37	31.7 X	146 bp
CRCH41-03 †	937.1 M	81.7%	37	32.3 X	147 bp
CRCH65-01	1024.7 M	85.2%	41	38.0 X	147 bp
CTAS71-01	816.4 M	98.1%	37	33.0 X	147 bp
CTAS72-04	992.4 M	99.3%	46	42.4 X	148 bp
CRVEEH77-02	880.7 M	98.6%	40	35.8 X	147 bp
CRVEEH78-01	986.9 M	98.3%	45	41.2 X	148 bp
CRVEEH79-02 †	1003.5 M	98.7%	46	41.3 X	147 bp
CQLD88-04	1016.0 M	90.5%	43	38.2 X	148 bp
CRCH90-02	967.5 M	51.8%	23	19.3 X	147 bp
CSA93-05	973.5 M	98.4%	45	40.7 X	147 bp
CSA100-01 †	964.5 M	98.6%	44	39.5 X	147 bp
CRVEEH113-01	961.8 M	97.4%	44	39.1 X	146 bp
CSA119-05 †	1019.0 M	98.3%	47	40.9 X	147 bp
CSA126-01	568.6 M	99.9%	45	33.7 X	241 bp
CSA128-02 †	942.1 M	97.7%	43	34.3 X	147 bp
CQLD130-01	1964.9 M	69.7%	65	47.5 X	147 bp
CRCH137-02	847.2 M	84.8%	33	29.8 X	148 bp
CSA152-01 †	985.2 M	99.4%	46	40.9 X	148 bp
CSA158-02 †	869.5 M	99.3%	40	37.2 X	148 bp
CSA167	1011.2 M	98.8%	47	42.3 X	148 bp
CSA168-01 †	874.0M	98.7%	40	35.9 X	147 bp
CSA169-02 †	933.9 M	99.7%	43	37.9 X	148 bp
CSA178-02 †	995.0 M	99.3%	46	42.0 X	148 bp
CSA179	569.1 M	99.8%	45	32.9 X	240 bp
CSA181-01 †	930.7 M	98.8%	43	38.6 X	148 bp
CSA182-02	946.6 M	98.7%	44	37.7 X	147 bp
CSA192	906.5 M	99.2%	42	39.2 X	148 bp

MultiQC (v1.9) metrics. M; million. ‡ Average sequencing depth at sites in VCF file. Boxed data points if <30. †family with a proband screened in (Javadiyan et al. 2017).

Supplementary data

Table S2 Paediatric cataract genes selected for screening

Gene	Locus	Inheritance	MIM	References	Additional ocular phenotypes	Other
AGK	7q34	AR	610345	(Aldahmesh et al. 2012a) †		Sengers syndrome (MIM: 212350)
BFSP1	20p12.1	AD/AR	603307	(Ramachandran et al. 2007)		
BFSP2	3q22.1	AD/AR	603212	(Conley et al. 2000) and (Jakobs et al. 2000)		
CHMP4B	20q11.22	AD	610897	(Shiels et al. 2007)		
COL4A1	13q34	AD	120130	(Xia et al. 2014) †		Small-vessel brain disease 1 with or without ocular anomalies (MIM: 175780)
COL4A2	13q34	AD	120090	(Ha et al. 2016)		Brain small vessel disease 2 (BSVD2, [MIM: 614483])
CRYAA	21q22.3	AD/AR	123580	(Litt et al. 1998)	Microcornea, iris coloboma, nystagmus, microphthalmia	
CRYAB	11q23.1	AD/AR	123590	(Berry et al. 2001) †		Myofibrillar myopathy (MFM2, [MIM: 608810])
CRYBA1	17q11.2	AD/AR	123610	(Kannabiran et al. 1998)	Nystagmus	
CRYBA2	2q35	AD	600836	(Reis et al. 2013)		
CRYBA4	22q12.1	AD/AR	123631	(Billingsley et al. 2006)	Microphthalmia, microcornea	
CRYBB1	22q12.1	AD/AR	600929	(Willoughby et al. 2005)	Microphthalmia, microcornea, nystagmus	
CRYBB2	22q11.23	AD	123620	(Litt et al. 1997)	Microphthalmia, microcornea, strabismus	
CRYBB3	22q11.23	AD/AR	123630	(Riazuddin et al. 2005)	Microcornea	
CRYGA	2q33.3	AD	123660	(Li et al. 2016)		
CRYGB	2q33.3	AD	123670	(AlFadhli et al. 2012)		

Supplementary data

CRYGC	2q33.3	AD	123680	(Heon et al. 1999)	Microcornea, glaucoma, microphthalmia, nystagmus	
CRYGD	2q33.3	AD	123690	(Stephan et al. 1999) and (Heon et al. 1999)	Nystagmus, microcornea	
CRYGS	3q27.3	AD	123730	(Sun et al. 2005)	Lens subluxation	
CTDP1	18q23	AR	604927	(Tzifi et al. 2011)		Congenital cataract, facial dysmorphism and neuropathy (CCFDN, [MIM: 604168])
CYP27A1	2q35	AR	606530	(Khan et al. 2015)		Cerebrotendinous xanthomatosis (MIM: 21370)
CYP51A1	7q21.2	AR	601637	(Aldahmesh et al. 2012b)		
DNMBP	10q24.2	AR	611282	(Ansar et al. 2018)	Nystagmus, amblyopia, exotropia	
EPHA2	1p36.13	AD/AR	176946	(Shiels et al. 2008)	Posterior lenticonus	
EYA1	8q13.3	AD	601653	(Azuma et al. 2000) †	Persistence of pupillary membrane, corneal opacity	Branchio-oto-renal syndrome with or without cataract (BOR1, [MIM:113650])
FBN1	15q21.1	Sporadic	134797	(Li et al. 2016) †		Marfan syndrome (MIM: 154700), Weill-Marchesani syndrome (WMS2 [MIM: 608328])
FOXE3	1p33	AD/AR	601094	(Semina et al. 2001)	Anterior segment dysgenesis 2 (ASGD2, [MIM: 610256])	
FTL	19q13.33	AD	134790	(Girelli et al. 1995) and (Beaumont et al. 1995)		Hyperferritinemia with or without cataract (HHCS, [MIM: 600886])
FYCO1	3p21.31	AR	607182	(Chen et al. 2011)		

Supplementary data

GALK1	17q25.1	AR	604313	(Stambolian et al. 1995)		Galactokinase deficiency with cataract (MIM: 230200)
GCNT2	6p24.3-p24.2	AR	600429	(Yu et al. 2001)	Nystagmus	Adult i blood group with cataract (MIM: 110800)
GJA3	13q12.11	AD	121015	(Mackay et al. 1999)		
GJA8	1q21.2	AD/AR	600897	(Shiels et al. 1998)	Microphthalmia, nystagmus, secondary glaucoma microcornea, corneal opacity, sclerocornea, coloboma	
HSF4	16q22.1	AD/AR	602438	(Bu et al. 2002)	Nystagmus	
IARS2	1q41	AR	612801	(Li et al. 2018) †		Cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss and skeletal dysplasia (CAGSSS, [MIM: 616007])
LEMD2	6p21.31	AR	616312	(Boone et al. 2016)		
LIM2	19q13.41	AR	154045	(Pras et al. 2002)	Nystagmus, amblyopia	
LONP1	19p13.3	AR	605490	(Khan et al. 2015)		CODAS syndrome (MIM: 600373)
LSS	21q22.3	AR	600909	(Zhao et al. 2015)		
MAF	16q23.2	AD	177075	(Jamieson et al. 2002) †	Microcornea, iris coloboma, amblyopia	Ayme-Gripp syndrome (MIM: 601088)
MIP	12q13.3	AD	154050	(Berry et al. 2000)	Nystagmus, strabismus	
MIR184	15q25.1	AD	613146	(Iliff et al. 2012)	EDICT syndrome (MIM: 614303)	
NHS	Xp22.2-p22.1	XL	300457	(Coccia et al. 2009) †	Microcornea, nystagmus, secondary glaucoma, strabismus	Nance-Horan syndrome (MIM: 302350)
PANK4	1p36.32	AD	606162	(Sun et al. 2019)	Nystagmus	

Supplementary data

PAX6	11p13	AD	607108	(Glaser et al. 1994)	Aniridia (MIM: 106210), nystagmus, corneal abnormalities, glaucoma	
PEX11B	1q21.1	AR	603867	(Taylor et al. 2017)		Peroxisome biogenesis disorder 14B (PBD14B, [MIM: 614920])
PITX3	10q24.32	AD/AR	602669	(Semina et al. 1998)	Anterior segment dysgenesis 1 (ASGD1, [MIM: 107250])	
PRX	19q13.2	AD	605725	(Yuan et al. 2016)	Amblyopia	
LOC105378949	1p36.33	AD	-	(Eiberg et al. 2019)		
RRAGA	9p22.1	AD	612194	(Chen et al. 2016)		
SLC16A12	10q23.31	AD	611910	(Kloeckener-Gruissem et al. 2008)	Microcornea	Glucosuria
SLC40A1	2q32.2	AD	604653	(Yamakawa et al. 2016)		Hemochromatosis (MIM: 606069)
SLC7A8	14q11.2	AR	604235	(Knopfel et al. 2019)		
SIPA1L3	19q13.1-q13.2	AR	616655	(Greenlees et al. 2015) and (Evers et al. 2015)	Corneal clouding, microphthalmia, iridocorneal and lenticular adhesions	
TDRD7	9q22.33	AR	611258	(Lachke et al. 2011)		
TMEM114	16p13.2	AD	611579	(Jamieson et al. 2007)		
TRPM3	9q21.12-q21.13	AD	608961	(Bennett et al. 2014)	Open-angle glaucoma	
UNC45B	17q12	AD	611220	(Hansen et al. 2014)		
VIM	10p13	AD	193060	(Muller et al. 2009)		
VSX2	14q24.3	AR	142993	(Percin et al. 2000)	Microphthalmia, iris coloboma, anophthalmia	
WDR36	5q22.1	Sporadic	609669	(Li et al. 2016)		
WDR87	19q13.13	AR	-	(Khan et al. 2015)		

Supplementary data

WFS1	4p16.1	AD	606201	(Berry et al. 2013) †	Iris coloboma	Wolfram syndrome 1 (WFS1, [MIM: 222300])

Inheritance; 'AD' autosomal dominant, 'AR' autosomal recessive, 'XL' X-linked, and 'sporadic' if only sporadic non-syndromic case has been reported in Cat-Map prior to publication (Shiels et al. 2010). MIM; gene reference. '-' data unavailable. Reference; initial gene reference. † reference for report of non-syndromic occurrence when gene is primarily known to cause syndromic congenital cataracts. Additional ocular phenotypes; as reported in Cat-Map (Shiels et al. 2010). Other; report of syndrome or condition associated with this gene.

Supplementary data

Table S3 Primers use for reported variant validation and segregation analysis

Family ID	Gene/variant	Primer pair (5' to 3')	Anneal temperature (°C)
CRCH4	*MIP* p.(Arg113Gln)	CCACCTGTCAATCCTCACCA TGTTCTGCAGGTGGCTATGG	57
CRCH5	*LONP1* p.(Pro553Ser)	AATGGGAATGGCTTTGGGGT TACAAGATTGTCAGCGGCGA	59
CRCH21	*GJA3* p.(Pro59Leu)	CTCTTCCATGCGCACGATGT GGAATCTGAAGCAATGGGCG	60
CRCH28	*EYA1* p.(Ser487Leu)	AATGCTGGGATGAGCTGAGTAG TAAATCCTCAGGTCTGCTTGG	57
	PITX3 p.(Ter303LeuextTer100)	TGAAAACGAGGGAGGGGAAG ACCCGTGTAACTCGAGCCT	61
CRCH29	*CRYAA* p.(Arg12Cys)	CTCCATTCTGCTGGTGGCA CAAGACCAGAGTCCATCGCT	59
CTAS34	*MIP* p.(Val164Ile)	CTACCTTGGGGTCAAGAAGGA CTTGAGGAGGTAACACTGTGGC	57
	CRYBB1 p.(Ile94Asn)	ATTTCTCCAGAGCCCAGAACCA GGATGGGAGGACAGGATCATT	57
CRCH38	*COL4A1* p.(Gly720Asp)	ATGTCCTGGGACGTTCACAA AAGTGGGGAACGGCATTGTA	57
	MIR184 n.52T>C	CCGGGAAATCAAACGTCCAT AACGCCAGGTTTTCCCCATC	57
CTAS71	*GJA8* p.(Lys131del)	CATGGAGGAGAAGCGCAAAAG GAAGTAGTGGCCCACGATGA	57
CRVEEH77	*GJA8* p.(Pro189Ser)	ACCCTGCTGAGGACCTACAT GACACAGAGGCCACAGACAA	57
CRVEEH79	*LEMD2* p.(?)	TTGTTGCGGTAGACATCCCG AAAGGCCAAGTGCAGACCTT	57
CQLD88	*FBN1* p.(Pro1141Leu)	GCTTCCAACTTTGGCAATGA GAGGCCCCCACCTTTAACAT	57
CRCH90	*GJA3* p.(Thr19Met)	TGTCGTAGCAGACGTTCTCG CCCGGTGTTCATGAGCATTT	57
CSA93	*PRX* p.(Arg129His)	AGGGGCAGAGGGTGGAATTA ATGCGCCGAGCCTTACAAAG	57
CSA100	*LSS* p.(Leu78Val)	AGTGGGCCACCATAATCACC TTGGGCTGTATGTGAAGAGGG	57
CQLD130	*BFSP2* p.(Arg89Trp)	CTCCAGGACCAATGCCATGAG TGTTTCCAGCTCCTGACTGAC	57
CRCH137	*GJA8* p.(Gly22Ser)	TCGGGGCCTTCTTTGTTCTC GCGAATGTGGGAGATGGGAA	57
CSA158	*WFS1* p.(Ala370Val)	CCCACGCACCACATCAAC CATAGGGCTCCAGGTGGTTC	63
	CYP51A1 p.(Ala94Thr)	ACCCCAGGACATGGGAAAAG GGTCATGAAACGAAACTGGG	60
CSA168	*HSF4* p.(Lys64Glu)	TGGTAGAGCGGGACCAGTTT CACCTTCCGAAAACCGTCTG	63
CSA178	*IARS2* p.(Gly389Ala)	CCCCACAGGTGTAGATTTGGA CAGTACCATGGGCAGGTTGT	57
CSA182	*BFSP1* p.(Glu375GlyfsTer2)	ATCCTCTGGAGCCCCTTCTT GCCTATTTTCCAACCAGCGT	57
CSA192	*GJA3* p.(Gln15Lys)	TGTCGTAGCAGACGTTCTCG CCCGGTGTTCATGAGCATTT	57

+Q; Optimized with Qiagen Q solution

Supplementary data

Table S4 Evidence for pathogenicity according to ACMG-AMP guidelines.

Family ID	Gene	Variant	Supporting evidence for pathogenicity	Supporting evidence for benign	ACMG-AMP	ClinVar
Pathogenic/Likely pathogenic						
CRCH21	*GJA3*	(NC_000013.10)Chr13:g.20717252G>A NM_021954.4:c.176C>T p.(Pro59Leu)	PM2, PP1_strong, PP3 Absent/near absent from population databases. Same variant previously established as pathogenic variant (report with cataract PMID: 15208569, 19182255, 21866213, 25148791, 26694549, 27609163), segregation (strong) with the disease in three families with 19 meioses (this study and PMID: 15208569, 27609163). Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV001573165
CRCH90	*GJA3*	(NC_000013.10)Chr13:g.20717372G>A NM_021954.4:c.56C>T p.(Thr19Met)	PM2, PP1_strong, PP3 Absent/near absent from population databases. Same variant has been reported as pathogenic (report with cataract PMID: 28839118, 21031021, 29461512), segregation (strong) with the disease in four families with five meioses (this study and PMID: 28839118, 21031021, 29461512). Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV001573166
CRCH29	*CRYAA*	(NC_000021.8)Chr21:g.44589243C>T NM_000394.4:c.34C>T p.(Arg12Cys)	PM2, PP1_strong, PP3 Absent/near absent from population databases. Same variant previously established pathogenic (report with cataract PMID: 23508780, 19503744, 17724170, 18587492, 19390652, 21686328, 30078984, 32010934), segregation (strong) with the disease in nine families with 20 meioses (this study and PMID:	-	LP	SCV001573167

Supplementary data

			23508780, 19503744, 17724170, 18587492, 19390652, 21686328, 30078984, 32010934). Multiple predictive tools assessing variant as damaging/pathogenic.			
CRCH38	*COL4A1*	(NC_000013.10)Chr13:g.110833673C>T NM_001845.6:c.2159G>A p.(Gly720Asp)	PM2, PP1_strong, PP3 Absent/near absent from population databases. Alternate source reports variant as pathogenic in family with additional congenital cataract phenotype (PMID: 17696175), segregation (strong) with the disease in two families with 6 meioses (this study and PMID: 17696175). Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV001573168
	MIR184	(NC_000015.9)Chr15:g.79502181T>C NR_029705.1:n.52T>C	PM2, PP3 Absent/near absent from population databases. Multiple predictive tools assessing variant as damaging/pathogenic.	BP4 RNA folding predictive evidence of non-damaging/benign effect on product.	VUS	SCV001573169
CSA168	*HSF4*	(NC_000016.9)Chr16:g.67199491A>G NM_001040667.3:c.190A>G p.(Lys64Glu)	PM1, PM2, PP1_strong, PP3 Variant located in functional protein region, a highly conserved DNA binding domain with multiple pathogenic variants. Absent/near absent from population databases. Segregation (strong) with the disease in one family previously reported with eight meioses (PMID: 29243736). Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV001573170
CRCH28	*PITX3*	(NC_000010.10)Chr10:g.103990272C>A NM_005029.4:c.908G>T p.(Ter303LeuextTer100)	PM2, PM4, PP1_supporting, PP3 Absent/near absent from population databases. Protein length altered due to loss of native stop codon. Supporting segregation evidence in the family with ≥3 meioses. Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV001573171

Supplementary data

	EYA1	(NC_000008.10)Chr8:g.72127864G>A NM_000503.6:c.1460C>T p.(Ser487Leu)	-	BS4, BP4, BS1 Lack of segregation in affected family members, multiple predictive tools assessing the variant as non-damaging/benign. Present in population databases more than expected for the condition.	B	SCV0015 73172
CSA182	*BFSP1*	(NC_000020.10)Chr20:g.17475593del NM_001195.5:c.1124del p.(Glu375GlyfsTer2)	PM2, PM4, PP1_strong, PP3 Absent/near absent from population databases. Protein length changes due to introduction of premature stop codon. PP1_strong segregation evidence with ≥5 meioses in family. Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV0015 73173
CRVEEH77	*GJA8*	(NC_000001.10)Chr1:g.147380647C>T NM_005267.5:c.565C>T p.(Pro189Ser)	PM2, PM5, PP1_supporting, PP3 Absent/near absent from population databases. Variant at an amino acid where a different missense variant was predicted pathogenic (PMID: 17724170). Segregation (supporting) with the disease in the family with ≥3 meioses. Multiple predictive tools assessing variant as damaging/pathogenic.	-	LP	SCV0015 73174
Variants of uncertain significance with evidence towards pathogenicity						
CRCH137	*GJA8*	(NC_000001.10)Chr1:g.147380146G>A NM_005267.5:c.64G>A p.(Gly22Ser)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV0015 73175
CTAS71	*GJA8*	(NC_000001.10)Chr1:g.147380470_147380472del NM_005267.5:c.388_390del p.(Lys131del)	PM2, PM4_supporting, PP3 Absent/near absent from population databases. In-frame deletion changes protein length, but in cytoplasmic region	-	VUS	SCV0015 73176

Supplementary data

			with variable amino acid conservation between species. Multiple predictive tools assessing variant as damaging/pathogenic.			
CSA192	*GJA3*	(NC_000013.10)Chr13:g.20717385G>T NM_021954.4:c.43C>A p.(Gln15Lys)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV001573177
CRVEEH79	*LEMD2*	(NC_000006.11)Chr6:g.33756893T>C NM_001348710.2:c.1A>G p.(?)	PM2, PP3 Absent/near absent from population databases. Multiple predictive tools assessing variant as damaging/pathogenic. Note: Variant causes loss of start codon with likely loss of function, but haploinsufficiency has not been shown to cause disease before.	-	VUS	SCV001573178
CSA93	*PRX*	(NC_000019.9)Chr19:g.40904522C>T NM_020956.2:c.386G>A p.(Arg129His)	PM2, PP1_supporting, PP3 Absent/near absent from population databases. Segregation (supporting) with the disease in the family with ≥3 meioses. Multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV001573179
CTAS34	*CRYBB1*	(NC_000022.10)Chr22:g.27008054A>T NM_001887.4:c.281T>A p.(Ile94Asn)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV001573180
	MIP	(NC_000012.11)Chr12:g.56847410C>T NM_012064.4:c.490G>A p.(Val164Ile)	PM2 Variant observed at a low rate in population data.	BP4 Multiple predictive tools assess amino acid change as non-damaging.	VUS	SCV001573181
CQLD130	*BFSP2*	(NC_000003.11)chr3:g.133119192C>T NM_003571.4:c.265C>T p.(Arg89Trp)	PM2, PP1_supporting, PP3 Supporting for absent/near absent from population databases. Segregation (supporting) with the disease in the family with ≥3 meioses. Multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV001573182

Supplementary data

CRCH4	*MIP*	(NC_000012.11)Chr12:g.56848060C>T NM_012064.4:c.338G>A p.(Arg113Gln)	PM2, PP3 Absent/near absent from population databases. Multiple predictive tools assessing variant as damaging/pathogenic.	BS2 Observed in health adult obligate carrier (and 2 other unaffected children) with full penetrance expected at an early age.	VUS	SCV001573183
Variants of uncertain significance that are unlikely to be causing disease						
CSA158	*CYP51A1*	(NC_000007.13)Chr7:g.91761099C>T NM_000786.4:c.280G>A p.(Ala94Thr)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	-	VUS	SCV001573184
	WFS1	(NC_000004.11)Chr4:g.6302631C>T NM_001145853.1:c.1109C>T p.(Ala370Val)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	BS4 Lack of segregation in an affected member of the family.	VUS	SCV001573185
CRCH5	*LONP1*	(NC_000019.9)Chr19:g.5694473G>A NM_001276480.1:c.1657C>T p.(Pro553Ser)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	BS4 Variant does not segregate with the disease in the family, absent from affected individual.	VUS	SCV001573186
CSA178	*IARS2*	(NC_000001.10)Chr1:g.220279332G>C NM_018060.4:c.1166G>C p.(Gly389Ala)	PM2, PP3 Absent/near absent from population databases and multiple predictive tools assessing variant as damaging/pathogenic.	BS4 Lack of segregation in affected member of family.	VUS	SCV001573187
CQLD88	*FBN1*	(NC_000015.9)Chr15:g.48779550G>A NM_000138.5:c.3422C>T p.(Pro1141Leu)	PP3 Multiple predictive tools assessing variant as damaging/pathogenic.	BS1, BS2 Present in the population database more than expected for the disease. Variant is observed in unaffected individuals in the family.	VUS	SCV001573188

Supplementary data

CSA100	*LSS*	(NC_000021.8)Chr21:g.47647553G>C NM_001001438.3:c.232C>G p.(Leu78Val)	PM2 Absent/near absent from population databases.	BS4, BP4 Lack of segregation with affected members of the family. Multiple predictive tools assess the amino acid change as non-pathogenic.	VUS	SCV001573189

Variant interpretation based on the American College of Medical Genetics and Genomics (ACMG-AMP) guidelines (Richards et al. 2015), 'P' Pathogenic, 'LP' Likely Pathogenic, 'VUS' Variant of Uncertain Significance, and 'B' Benign. ClinVar; variant accession number.

Supplementary data

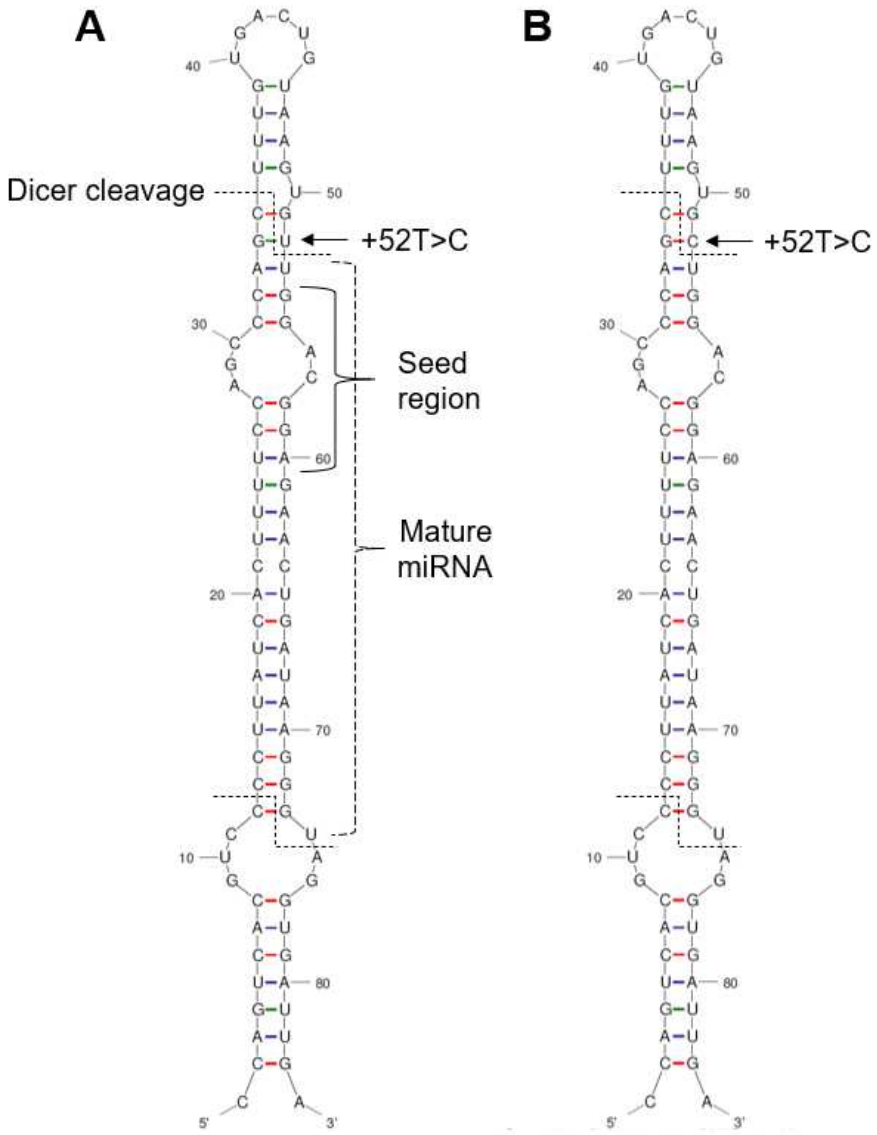

Figure S1 Predicted impact of the NR_029705.1:n.+52T>C variant on miR-184 structure and stability. The *MIR184* variant in family CRCH38 was assessed using mFold (Zuker 2003). No difference in miRNA folding was observed between wild-type miR-184 (**A**) and +52T>C variant carrier miR-184 (**B**). Minimum free energy prediction dropped from -35.50kcal/mol to -37.90kcal/mol with the T>C nucleotide change, increasing the stability of the molecule.

Supplementary data

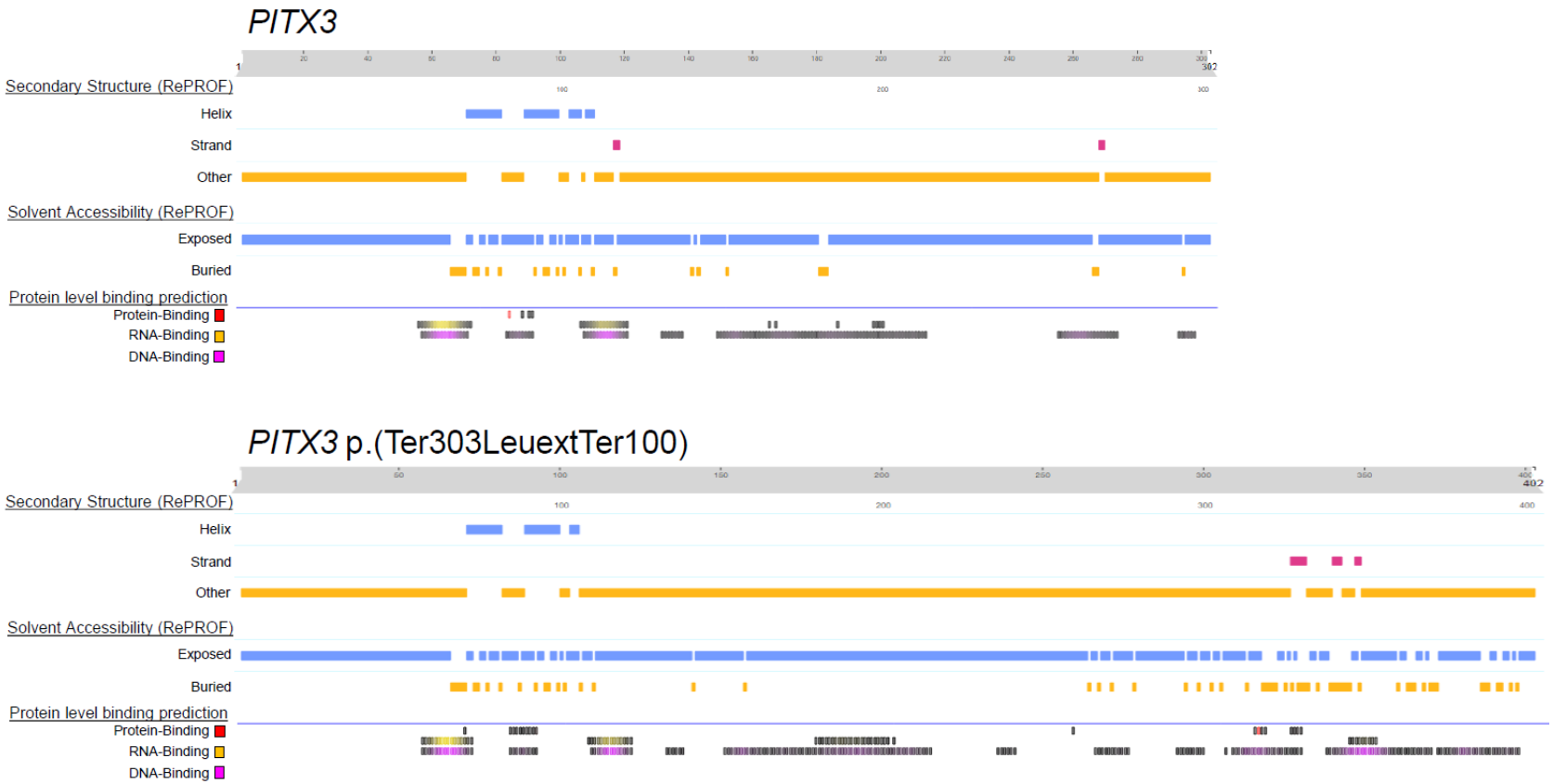

Figure S2 *PITX3* protein folding prediction. Comparison of wild-type *PITX3* and CRCH28 family stop loss p.(Ter303LeuextTer100) *PITX3* variant sequence using PredictProtein folding prediction tool (https://predictprotein.org). Each panel displays predicted secondary structure and solvent accessibility, with protein level binding predictions appended below. *PITX3*_variant sequence has been stretched to enable vertical comparison of the complementary initial 1-302 amino acid positions.

Supplementary data

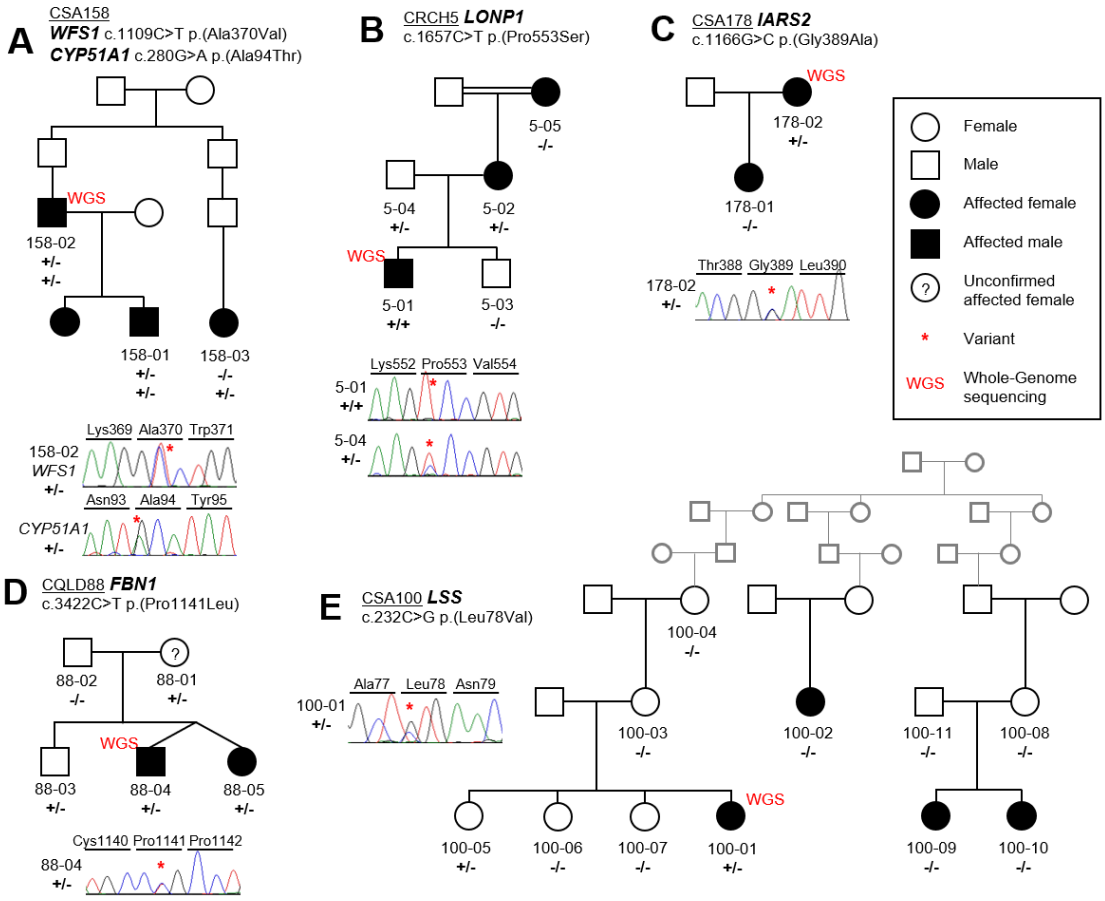

Figure S3 Variants of uncertain significance that are unlikely to be disease causing. Family pedigrees and accompanying sequencing chromatograms from variant analysis. A; Family CSA158 with a segregating *CYP51A1* p.(Ala94Thr) and non-segregating *WFS1* p.(Ala370Val) variants. B; *LONP1* variant enters pedigree twice to enable homozygous status in the proband in family CRCH5. C; In family CSA178, an *IARS2* variant was observed affected mother only. D; In family CQLD88, a *FBN1* variant was observed in an unaffected individual. E; Large pedigree CSA100 displays a *LSS* variant in an affected and unaffected sibling pair, that has entered in the pedigree in the previous generation.

Supplementary data

References:

Aldahmesh MA, Khan AO, Mohamed JY, Alghamdi MH, Alkuraya FS. 2012a. Identification of a truncation mutation of acylglycerol kinase (agk) gene in a novel autosomal recessive cataract locus. Human mutation. 33(6):960-962.

Aldahmesh MA, Khan AO, Mohamed JY, Hijazi H, Al-Owain M, Alswaid A, Alkuraya FS. 2012b. Genomic analysis of pediatric cataract in saudi arabia reveals novel candidate disease genes. Genetics in medicine : official journal of the American College of Medical Genetics. 14(12):955-962.

AlFadhli S, Abdelmoaty S, Al-Hajeri A, Behbehani A, Alkuraya F. 2012. Novel crystallin gamma b mutations in a kuwaiti family with autosomal dominant congenital cataracts reveal genetic and clinical heterogeneity. Molecular vision. 18:2931-2936.

Ansar M, Chung HL, Taylor RL, Nazir A, Imtiaz S, Sarwar MT, Manousopoulou A, Makrythanasis P, Saeed S, Falconnet E et al. 2018. Bi-allelic loss-of-function variants in dnmbp cause infantile cataracts. American journal of human genetics. 103(4):568-578.

Azuma N, Hirakiyama A, Inoue T, Asaka A, Yamada M. 2000. Mutations of a human homologue of the drosophila eyes absent gene (eya1) detected in patients with congenital cataracts and ocular anterior segment anomalies. Human molecular genetics. 9(3):363-366.

Beaumont C, Leneuve P, Devaux I, Scoazec JY, Berthier M, Loiseau MN, Grandchamp B, Bonneau D. 1995. Mutation in the iron responsive element of the l ferritin mrna in a family with dominant hyperferritinaemia and cataract. Nature genetics. 11(4):444-446.

Bennett TM, Mackay DS, Siegfried CJ, Shiels A. 2014. Mutation of the melastatin-related cation channel, trpm3, underlies inherited cataract and glaucoma. PloS one. 9(8):e104000.

Berry V, Francis P, Kaushal S, Moore A, Bhattacharya S. 2000. Missense mutations in mip underlie autosomal dominant 'polymorphic' and lamellar cataracts linked to 12q. Nature genetics. 25(1):15-17.

Berry V, Francis P, Reddy MA, Collyer D, Vithana E, MacKay I, Dawson G, Carey AH, Moore A, Bhattacharya SS et al. 2001. Alpha-b crystallin gene (cryab)

mutation causes dominant congenital posterior polar cataract in humans. American journal of human genetics. 69(5):1141-1145.

Berry V, Gregory-Evans C Fau - Emmett W, Emmett W Fau - Waseem N, Waseem N Fau - Raby J, Raby J Fau - Prescott D, Prescott D Fau - Moore AT, Moore At Fau - Bhattacharya SS, Bhattacharya SS. 2013. Wolfram gene (wfs1) mutation causes autosomal dominant congenital nuclear cataract in humans. (1476-5438 (Electronic)).

Billingsley G, Santhiya ST, Paterson AD, Ogata K, Wodak S, Hosseini SM, Manisastry SM, Vijayalakshmi P, Gopinath PM, Graw J et al. 2006. Cryba4, a novel human cataract gene, is also involved in microphthalmia. American journal of human genetics. 79(4):702-709.

Boone PM, Yuan B, Gu S, Ma Z, Gambin T, Gonzaga-Jauregui C, Jain M, Murdock TJ, White JJ, Jhangiani SN et al. 2016. Hutterite-type cataract maps to chromosome 6p21.32-p21.31, cosegregates with a homozygous mutation in lemd2, and is associated with sudden cardiac death. Molecular genetics & genomic medicine. 4(1):77-94.

Bu L, Jin Y, Shi Y, Chu R, Ban A, Eiberg H, Andres L, Jiang H, Zheng G, Qian M et al. 2002. Mutant DNA-binding domain of hsf4 is associated with autosomal dominant lamellar and marner cataract. Nature genetics. 31(3):276-278.

Chen J, Ma Z, Jiao X, Fariss R, Kantorow WL, Kantorow M, Pras E, Frydman M, Pras E, Riazuddin S et al. 2011. Mutations in fyco1 cause autosomal-recessive congenital cataracts. American journal of human genetics. 88(6):827-838.

Chen JH, Huang C, Zhang B, Yin S, Liang J, Xu C, Huang Y, Cen LP, Ng TK, Zheng C et al. 2016. Mutations of raga gtpase in mtorc1 pathway are associated with autosomal dominant cataracts. PLoS genetics. 12(6):e1006090.

Coccia M, Brooks SP, Webb TR, Christodoulou K, Wozniak IO, Murday V, Balicki M, Yee HA, Wangensteen T, Riise R et al. 2009. X-linked cataract and nance-horan syndrome are allelic disorders. Human molecular genetics. 18(14):2643-2655.

Conley YP, Erturk D, Keverline A, Mah TS, Keravala A, Barnes LR, Bruchis A, Hess JF, FitzGerald PG, Weeks DE et al. 2000. A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense

mutation in the beaded filament structural protein-2. American journal of human genetics. 66(4):1426-1431.

Eiberg H, Mikkelsen AF, Bak M, Tommerup N, Lund AM, Wenzel A, Sabarinathan R, Gorodkin J, Bang-Berthelsen CH, Hansen L. 2019. A splice-site variant in the lncrna gene rp1-140a9.1 cosegregates in the large volkmann cataract family. Molecular vision. 25:1-11.

Evers C, Paramasivam N, Hinderhofer K, Fischer C, Granzow M, Schmidt-Bacher A, Eils R, Steinbeisser H, Schlesner M, Moog U. 2015. Sipa1l3 identified by linkage analysis and whole-exome sequencing as a novel gene for autosomal recessive congenital cataract. European journal of human genetics : EJHG. 23(12):1627-1633.

Girelli D, Corrocher R, Bisceglia L, Olivieri O, De Franceschi L, Zelante L, Gasparini P. 1995. Molecular basis for the recently described hereditary hyperferritinemia-cataract syndrome: A mutation in the iron-responsive element of ferritin l-subunit gene (the "verona mutation"). Blood. 86(11):4050-4053.

Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL. 1994. Pax6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. Nature genetics. 7(4):463-471.

Greenlees R, Mihelec M, Yousoof S, Speidel D, Wu SK, Rinkwitz S, Prokudin I, Perveen R, Cheng A, Ma A et al. 2015. Mutations in sipa1l3 cause eye defects through disruption of cell polarity and cytoskeleton organization. Human molecular genetics. 24(20):5789-5804.

Ha TT, Sadleir LG, Mandelstam SA, Paterson SJ, Scheffer IE, Gecz J, Corbett MA. 2016. A mutation in col4a2 causes autosomal dominant porencephaly with cataracts. American journal of medical genetics Part A. 170a(4):1059-1063.

Hansen L, Comyn S, Mang Y, Lind-Thomsen A, Myhre L, Jean F, Eiberg H, Tommerup N, Rosenberg T, Pilgrim D. 2014. The myosin chaperone unc45b is involved in lens development and autosomal dominant juvenile cataract. European journal of human genetics : EJHG. 22(11):1290-1297.

Heon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, Lubsen N, Munier FL. 1999. The gamma-crystallins and human cataracts: A puzzle made clearer. American journal of human genetics. 65(5):1261-1267.

Supplementary data

Iliff BW, Riazuddin SA, Gottsch JD. 2012. A single-base substitution in the seed region of mir-184 causes edict syndrome. Investigative ophthalmology & visual science. 53(1):348-353.

Jakobs PM, Hess JF, FitzGerald PG, Kramer P, Weleber RG, Litt M. 2000. Autosomal-dominant congenital cataract associated with a deletion mutation in the human beaded filament protein gene bfsp2. American journal of human genetics. 66(4):1432-1436.

Jamieson RV, Farrar N, Stewart K, Perveen R, Mihelec M, Carette M, Grigg JR, McAvoy JW, Lovicu FJ, Tam PP et al. 2007. Characterization of a familial t(16;22) balanced translocation associated with congenital cataract leads to identification of a novel gene, tmem114, expressed in the lens and disrupted by the translocation. Human mutation. 28(10):968-977.

Jamieson RV, Perveen R, Kerr B, Carette M, Yardley J, Heon E, Wirth MG, van Heyningen V, Donnai D, Munier F et al. 2002. Domain disruption and mutation of the bzip transcription factor, maf, associated with cataract, ocular anterior segment dysgenesis and coloboma. Human molecular genetics. 11(1):33-42.

Javadiyan S, Craig JE, Souzeau E, Sharma S, Lower KM, Mackey DA, Staffieri SE, Elder JE, Taranath D, Straga T et al. 2017. High-throughput genetic screening of 51 pediatric cataract genes identifies causative mutations in inherited pediatric cataract in south eastern australia. G3 (Bethesda, Md). 7(10):3257-3268.

Kannabiran C, Rogan PK, Olmos L, Basti S, Rao GN, Kaiser-Kupfer M, Hejtmancik JF. 1998. Autosomal dominant zonular cataract with sutural opacities is associated with a splice mutation in the betaa3/a1-crystallin gene. Molecular vision. 4:21.

Khan AO, Aldahmesh MA, Alkuraya FS. 2015. Phenotypes of recessive pediatric cataract in a cohort of children with identified homozygous gene mutations (an american ophthalmological society thesis). Transactions of the American Ophthalmological Society. 113:T7.

Kloeckener-Gruissem B, Vandekerckhove K, Nürnberg G, Neidhardt J, Zeitz C, Nürnberg P, Schipper I, Berger W. 2008. Mutation of solute carrier slc16a12 associates with a syndrome combining juvenile cataract with microcornea and renal glucosuria. American journal of human genetics. 82(3):772-779.

Supplementary data

Knopfel EB, Vilches C, Camargo SMR, Errasti-Murugarren E, Staubli A, Mayayo C, Munier FL, Miroshnikova N, Poncet N, Junza A et al. 2019. Dysfunctional lat2 amino acid transporter is associated with cataract in mouse and humans. Frontiers in physiology. 10:688.

Lachke SA, Alkuraya FS, Kneeland SC, Ohn T, Aboukhalil A, Howell GR, Saadi I, Cavallesco R, Yue Y, Tsai AC et al. 2011. Mutations in the rna granule component tdrd7 cause cataract and glaucoma. Science (New York, NY). 331(6024):1571-1576.

Li D, Wang S, Ye H, Tang Y, Qiu X, Fan Q, Rong X, Liu X, Chen Y, Yang J et al. 2016. Distribution of gene mutations in sporadic congenital cataract in a han chinese population. Molecular vision. 22:589-598.

Li J, Leng Y, Han S, Yan L, Lu C, Luo Y, Zhang X, Cao L. 2018. Clinical and genetic characteristics of chinese patients with familial or sporadic pediatric cataract. Orphanet journal of rare diseases. 13(1):94.

Litt M, Carrero-Valenzuela R, LaMorticella DM, Schultz DW, Mitchell TN, Kramer P, Maumenee IH. 1997. Autosomal dominant cerulean cataract is associated with a chain termination mutation in the human beta-crystallin gene crybb2. Human molecular genetics. 6(5):665-668.

Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, Weleber RG. 1998. Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene cryaa. Human molecular genetics. 7(3):471-474.

Mackay D, Ionides A, Kibar Z, Rouleau G, Berry V, Moore A, Shiels A, Bhattacharya S. 1999. Connexin46 mutations in autosomal dominant congenital cataract. American journal of human genetics. 64(5):1357-1364.

Muller M, Bhattacharya SS, Moore T, Prescott Q, Wedig T, Herrmann H, Magin TM. 2009. Dominant cataract formation in association with a vimentin assembly disrupting mutation. Human molecular genetics. 18(6):1052-1057.

Percin EF, Ploder LA, Jessica JY, Arici K, Horsford DJ, Rutherford A, Bapat B, Cox DW, Duncan AM, Kalnins VI. 2000. Human microphthalmia associated with mutations in the retinal homeobox gene chx10. Nature genetics. 25(4):397-401.

Pras E, Levy-Nissenbaum E Fau - Bakhan T, Bakhan T Fau - Lahat H, Lahat H Fau - Assia E, Assia E Fau - Geffen-Carmi N, Geffen-Carmi N Fau - Frydman M,

Supplementary data

Frydman M Fau - Goldman B, Goldman B Fau - Pras E, Pras E. 2002. A missense mutation in the lim2 gene is associated with autosomal recessive presenile cataract in an inbred iraqi jewish family. (0002-9297 (Print)).

Ramachandran RD, Perumalsamy V, Hejtmancik JF. 2007. Autosomal recessive juvenile onset cataract associated with mutation in bfsp1. Human genetics. 121(3-4):475-482.

Reis LM, Tyler RC, Muheisen S, Raggio V, Salviati L, Han DP, Costakos D, Yonath H, Hall S, Power P et al. 2013. Whole exome sequencing in dominant cataract identifies a new causative factor, cryba2, and a variety of novel alleles in known genes. Human genetics. 132(7):761-770.

Riazuddin SA, Yasmeen A, Yao W, Sergeev YV, Zhang Q, Zulfiqar F, Riaz A, Riazuddin S, Hejtmancik JF. 2005. Mutations in betab3-crystallin associated with autosomal recessive cataract in two pakistani families. Investigative ophthalmology & visual science. 46(6):2100-2106.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E. 2015. 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. Genetics in medicine. 17(5):405-423.

Semina EV, Brownell I, Mintz-Hittner HA, Murray JC, Jamrich M. 2001. Mutations in the human forkhead transcription factor foxe3 associated with anterior segment ocular dysgenesis and cataracts. Human molecular genetics. 10(3):231-236.

Semina EV, Ferrell RE, Mintz-Hittner HA, Bitoun P, Alward WL, Reiter RS, Funkhauser C, Daack-Hirsch S, Murray JC. 1998. A novel homeobox gene pitx3 is mutated in families with autosomal-dominant cataracts and asmd. Nature genetics. 19(2):167-170.

Shiels A, Bennett TM, Hejtmancik JF. 2010. Cat-map: Putting cataract on the map. Molecular vision. 16:2007-2015.

Shiels A, Bennett TM, Knopf HL, Maraini G, Li A, Jiao X, Hejtmancik JF. 2008. The epha2 gene is associated with cataracts linked to chromosome 1p. Molecular vision. 14:2042-2055.

Shiels A, Bennett TM, Knopf HL, Yamada K, Yoshiura K, Niikawa N, Shim S, Hanson PI. 2007. Chmp4b, a novel gene for autosomal dominant cataracts

linked to chromosome 20q. American journal of human genetics. 81(3):596-606.

Shiels A, Mackay D, Ionides A, Berry V, Moore A, Bhattacharya S. 1998. A missense mutation in the human connexin50 gene (gja8) underlies autosomal dominant "zonular pulverulent" cataract, on chromosome 1q. American journal of human genetics. 62(3):526-532.

Stambolian D, Ai Y, Sidjanin D, Nesburn K, Sathe G, Rosenberg M, Bergsma DJ. 1995. Cloning of the galactokinase cdna and identification of mutations in two families with cataracts. Nature genetics. 10(3):307-312.

Stephan DA, Gillanders E, Vanderveen D, Freas-Lutz D, Wistow G, Baxevanis AD, Robbins CM, VanAuken A, Quesenberry MI, Bailey-Wilson J et al. 1999. Progressive juvenile-onset punctate cataracts caused by mutation of the gammad-crystallin gene. Proceedings of the National Academy of Sciences of the United States of America. 96(3):1008-1012.

Sun H, Ma Z, Li Y, Liu B, Li Z, Ding X, Gao Y, Ma W, Tang X, Li X et al. 2005. Gamma-s crystallin gene (crygs) mutation causes dominant progressive cortical cataract in humans. Journal of medical genetics. 42(9):706-710.

Sun M, Chen C, Hou S, Li X, Wang H, Zhou J, Chen X, Liu P, Kijlstra A, Lin S et al. 2019. A novel mutation of pank4 causes autosomal dominant congenital posterior cataract. Human mutation. 40(4):380-391.

Taylor RL, Handley MT, Waller S, Campbell C, Urquhart J, Meynert AM, Ellingford JM, Donnelly D, Wilcox G, Lloyd IC et al. 2017. Novel pex11b mutations extend the peroxisome biogenesis disorder 14b phenotypic spectrum and underscore congenital cataract as an early feature. Investigative ophthalmology & visual science. 58(1):594-603.

Tzifi F, Pons R, Athanassaki C, Poulou M, Kanavakis E. 2011. Congenital cataracts, facial dysmorphism, and neuropathy syndrome. Pediatric neurology. 45(3):206-208.

Willoughby CE, Shafiq A, Ferrini W, Chan LL, Billingsley G, Priston M, Mok C, Chandna A, Kaye S, Heon E. 2005. Crybb1 mutation associated with congenital cataract and microcornea. Molecular vision. 11:587-593.

Xia XY, Li N, Cao X, Wu QY, Li TF, Zhang C, Li WW, Cui YX, Li XJ, Xue CY. 2014. A novel col4a1 gene mutation results in autosomal dominant non-syndromic congenital cataract in a chinese family. BMC medical genetics. 15:97.

Supplementary data

Yamakawa N, Oe K, Yukawa N, Murakami K, Nakashima R, Imura Y, Yoshifuji H, Ohmura K, Miura Y, Tomosugi N et al. 2016. A novel phenotype of a hereditary hemochromatosis type 4 with ferroportin-1 mutation, presenting with juvenile cataracts. Internal medicine (Tokyo, Japan). 55(18):2697-2701.

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(13):3840-3845.

Yuan L, Yi J, Lin Q, Xu H, Deng X, Xiong W, Xiao J, Jiang C, Yuan X, Chen Y et al. 2016. Identification of a prx variant in a chinese family with congenital cataract by exome sequencing. QJM : monthly journal of the Association of Physicians. 109(11):731-735.

Zhao L, Chen XJ, Zhu J, Xi YB, Yang X, Hu LD, Ouyang H, Patel SH, Jin X, Lin D et al. 2015. Lanosterol reverses protein aggregation in cataracts. Nature. 523(7562):607-611.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic acids research. 31(13):3406-3415.