SUPPORTING INFORMATION

Methods and Results

Clinical assessment

Patients with a provisional diagnosis of arRP were collected and clinically examined in the Clinical Investigating Centre of the Quinze-Vingts Hospital. Informed consent was obtained from each patient and normal controls after explanation of the study and its potential outcome. The study protocol adhered to the tenets of the Declaration of Helsinki and was approved by the local ethics committee. Each patient underwent full ophthalmic examination with clinical assessment as described earlier. (Audo, et al., 2010). For additional family members who could not come to our centre for examination, ophthalmic records were obtained from local ophthalmologists.

Mutation detection by arRP microarray

Total genomic DNA was extracted from peripheral blood leucocytes according to manufacturer's recommendation (Puregen Kit, Qiagen, Courtaboeuf, France). The DNAs of 400 index patients were analyzed for known mutations by microarray analysis on a commercially available chip (arRP, ASPER Ophthalmics, Tartu, Estonia). Mutations identified by this approach were validated by direct Sanger sequencing. In cases where only one heterozygous mutation was detected, the second mutation was identified by direct sequencing of all exons and flanking intronic sequences of *CRB1* (NM_201253.2; including alternative transcript AF154671.1).

Out of 400 index patients nine probands were found to have *CRB1* mutations on the microarray. Two patients were homozygous and two other compound heterozygous for known mutations. Four patients were heterozygous for one known mutation and one patient showed an unexpected event in exon 6 of *CRB1*. Direct sequencing of this exon identified a novel frameshift mutation (p. Leu655Trpfs*10,) in a heterozygous state. All mutations identified by microarray analysis were confirmed by direct sequencing and the second mutation was identified in four of the five patients (Table 1 main text). Using this strategy we identified five novel *CRB1* mutations, two missense changes (p.Tyr1198Cys and p.Cys1223Ser), one nonsense mutation (p.Cys423*), one in-frame deletion (p.Asn789del) and one frameshift deletion mentioned above (p. Leu655Trpfs*10) (Table 1).

Homozygosity mapping

One consanguineous family (F709), excluded for known mutations by the first screening approach, was analysed using a 700K SNP microarray (HumanOmniExpress, Illumina, Eindhoven, The Netherlands). The SNP genotypes were analysed using commercially available software (GenomeStudio, Illumina, Eindhoven, The Netherlands) according to the protocols provided by Illumina. In the initial analysis, 686389 SNPs passed quality control. The homozygous regions were found through a web-based tool HomozygosityMapper (http://www.homozygositymapper.org/) (Seelow, et al., 2009).

The analysis revealed eight significant homozygous regions on chromosome 1 (16, 17 and 53 Mb), chromosome 4 (29 Mb), chromosome 6 (16 and 20 Mb) and chromosome 12 (13 and 56 Mb). These homozygous regions contained ten known retinopathy genes: (*ABCA4, PRPF3, SEMA4A, CRB1, CC2D2A, BBS7, BBS12, PROM1, BBS10, CEP290*) of which *CRB1*

was the most promising candidate as suggested by the patient's phenotype. *CRB1* was located in a 17 Mb homozygous region on chromosome 1, which was the 4th largest homozygous region. Direct sequencing of *CRB1* revealed a novel homozygous deletion-insertion in exon 9 (c.3659_3660delinsA, p.Ser1220Asnfs*62) (Table 1).

Next generation sequencing (NGS)

One consanguineous family was investigated by NGS using a custom-made oligonucleotide library targeting 177 known genes underlying retinal disorders (http://www.sph.uth.tmc.edu/retnet/sum-dis.htm, October 2010) and additional candidate genes (Audo et al., 2011 "Application of next-generation-sequencing (NGS) allows novel genotype-phenotype correlations of retinal diseases"). A custom-made SureSelect oligonucleotide probe library was designed to capture the exons according to Agilent's recommendations, using the eArray web-based probe design tool (https://earray.chem.agilent.com/earray). The following parameters were chosen for probe design: 120 bp length, 3x probe-tiling frequency, 20 bp overlap allowed in avoided region and exclusion of repetitive DNA sequences identified by implementing eArray's RepeatMasker program. A total of 27 430 probes, covering 1 177 Mb, were designed and synthesized by Agilent Technologies (Santa Clara, CA, USA). Sequence capture, enrichment, and elution were performed according to the manufacturer's instructions (SureSelect, Agilent). Briefly, 3 µg of each genomic DNA were fragmented by sonication and purified to yield fragments of 150-200 bp. Paired-end adaptor oligonucleotides from Illumina were ligated on repaired DNA fragments, which were then purified and enriched by 6 PCR cycles. 500ng of the purified libraries were hybridized to the SureSelect oligo probe capture library for 24h. After hybridization, washing, and elution, the eluted fraction was PCR-amplified with 14 cycles, purified and quantified by qPCR to obtain sufficient DNA template for downstream applications. Each eluted-enriched DNA sample was then sequenced on an Illumina GAIIx as paired-end 75 bp reads. Image analysis and base calling was performed using Illumina Real Time Analysis (RTA) Pipeline version 1.10 with default parameters. Sequence reads were aligned to the reference human genome (UCSC hg19) using commercially available software (CASAVA1.7, Illumina) and the ELANDv2 alignment algorithm. Genetics variation annotation was performed using the in-house pipeline, which consisted of gene annotation (RefSeq), detection of known polymorphisms (dbSNP 131, 1000 Genome) followed by a mutation characterization (exonic, intronic, silent, nonsense etc.). For each position, the exomic frequencies (homozygous and heterozygous) were determined from all the exomes already sequenced by Integragen, and the exome results provided by HapMap project.

The first screening criteria applied to the index patient form the consanguineous family were absence of the variant in dbSNP databases and homozygous appearance. This initial screen resulted in three homozygous mutations, of which p.Ser740Phe exchange in *CRB1* was the most convincing (Table 1 in the main test). This mutation was confirmed by Sanger sequencing and by performing cosegregation analysis in the family members (Figure 1). More details on data analysis from the NGS study of retinal genes are published elsewhere (Audo et al., 2011 "Application of next-generation-sequencing (NGS) allows novel genotype-phenotype correlations of retinal diseases").

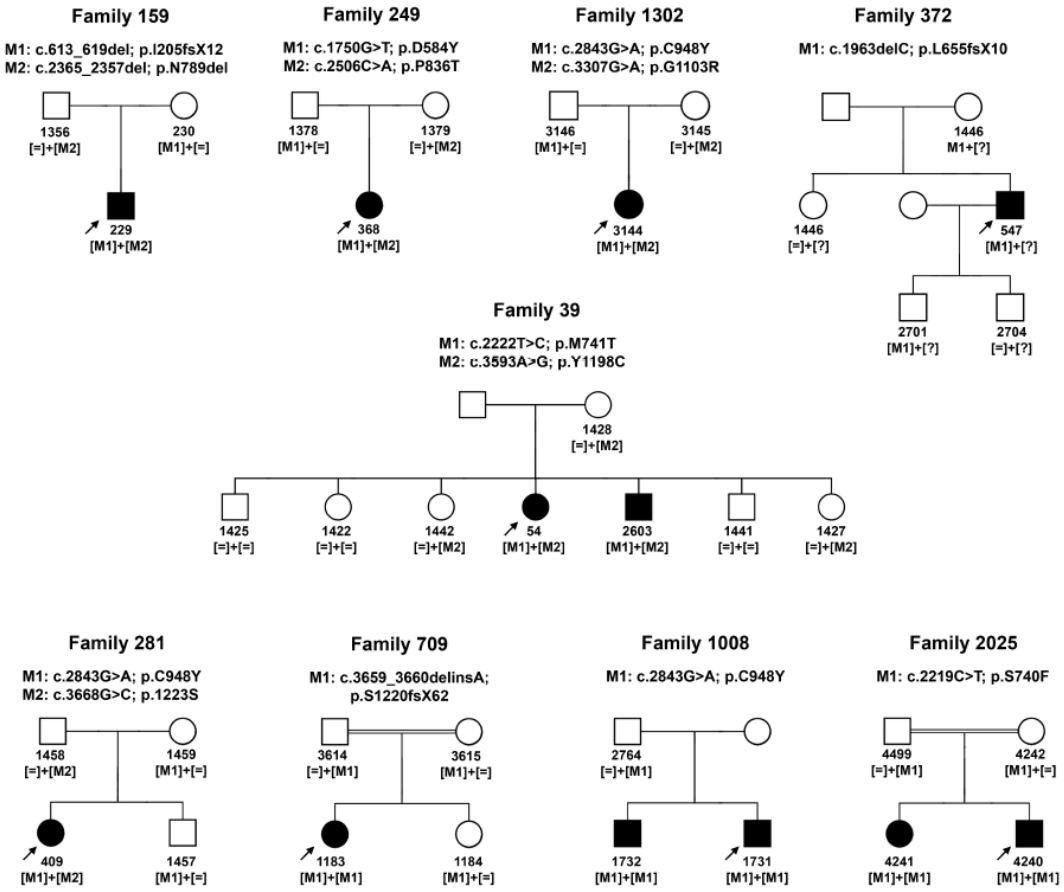
Sanger sequencing

For Sanger sequencing, *CRB1* gene (*CRB1* RefSeq NM_201253) was PCR amplified in 15 fragments using oligonucl eotides flanking the exons and a polymerase (HotFire, Solis

Biodyne, Estonia) in the presence of 1.5-2.0 mM $MgCl_2$ and at an annealing temperature of 55°C. The PCR products were enzymatically purified (ExoSAP-IT, USB Corporation, Cleveland, Ohio, USA purchased from GE Healthcare, Orsay, France) and sequenced with a commercially available sequencing mix (BigDyeTerm v1.1 CycleSeq kit, Applied Biosystems, Courtaboeuf, France). The sequenced products were purified on a presoaked Sephadex G-50 (GE Healthcare) 96-well multiscreen filter plate (Millipore, Molsheim, France), the purified product analyzed on an automated 48-capillary sequencer (ABI 3730 Genetic analyzer, Applied Biosystems) and the results interpreted by applying a software (SeqScape, Applied Biosystems). At least 362 commercially available control chromosomes were used to validate the pathogenicity of the novel sequence variants (Human random control panel 1-3, Health Protection Agency Culture Collections, Salisbury, United Kingdom).

Mutation nomenclature and assessment of the pathogenicity of mutations

Nucleotide numbering is based on cDNA sequence of *CRB1* (Ref. NM_201253.2) where A of the ATG initiation codon is 1. To evaluate the pathogenicity of the novel changes we applied the following criteria: 1) stop/frameshift mutations are most likely disease causing; 2) cosegregation in the family; 3) absence in control samples; 4) for missense mutations and in-frame deletions, amino acid conservation was studied in the UCSC Genome Browser using 27 species belonging to different evolutionary branches (Human, Chimp, Gorilla, Rhesus, Tarsier, Mouse lemur, Bushbaby, Tree shrew, Mouse, Squirrel, Rabbit, Cow, Horse, Cat, Dog, Hedgehog, Elephant, Sloth, Wallaby, Opossum, Platypus, Chicken, Lizard, X.tropicalis, Tetraodon, Stickleback and Zebrafish); if the amino acid residue did not change throughout the species it was considered as “highly conserved”; if a change was seen in fewer than five species and not in the primates then it was considered as “moderately conserved”; if a change was present in 5-7, it was considered as “weakly conserved”; otherwise the amino acid residue was considered as “not conserved”; 5) pathogenicity predictions with bioinformatic tools (PolyPhen-2, Polymorphism Phenotyping, http://genetics.bwh.harvard.edu/pph2/ (Adzhubei, et al.), and SIFT, Sorting Intolerant From Tolerant; http://blocks.fhcrc.org/sift/SIFT.html (Ng and Henikoff, 2003)); 6) presence of the second mutant allele. These criteria were applied to the mutations found in the patients described in this study as well as for the previously published mutations. All the variants were classified into three groups: likely pathogenic; unclassified variants, unlikely pathogenic. This classification is only indicative and has been based on the above criteria.



Supp. Figure S1. Cosegregation analysis of *CRB1* mutations in nine arRP families. Circles indicate females and squares males, the filled symbols represent affected individuals and the empty symbols denote healthy family members. Arrows indicate index patients and the question mark denotes an unknown allele. Cosegregation in patients 53 and 3969 is not represented due to unavailable family members.

Supp. Table S1. Likely pathogenic mutations in *CRB1*

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
2	c.107C>G	p.Ser36*	EGF1	protein truncation, NMD	-	-	2	LCA		(McKibbin, et al., 2010)
2	c.111delT	p.Ser38Leufs*33	EGF1	protein truncation, NMD	-	-	1	LCA	unknown second allele	(Lotery, et al., 2001a)
2	c.135C>G	p.Cys45Trp	EGF1	Highly conserved (considering 23 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.997)	1	RP	unknown second allele	(Clark, et al.)
2	c.257_258dupTG	p.Asn87*	EGF2	protein truncation, NMD	-	-	2	LCA		(Jacobson, et al., 2003; Lotery, et al., 2001a)
2	c.258C>T	p.Gln120*	EGF3	protein truncation, NMD	-	-	2	LCA		(Simonelli, et al., 2007)
2	c.428_432delGATTC	p.Arg143Metfs*2	EGF3	protein truncation, NMD	-	-	1	LCA	unknown second allele	(Lotery, et al., 2001a)
2	c.430T>G	p.Phe144Val	EGF3	Highly conserved in placental mammals (considering 18 species)	Tolerated (score 0.50)	Possibly Damaging (score 0.600)	1	LCA	unknown second allele	(Lotery, et al., 2001a)
2	c.470G>C	p.Cys157Ser	EGF4	Highly conserved (considering 26 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.996)	1	EOCRD		(Henderson, et al., 2010)
2	c.481dupG	p.Ala161Glyfs*8	EGF4	protein truncation, NMD	-	-	5	RP, LCA, EORP,		(Bernal, et al., 2003; Vallespin, et al., 2007b)
2	c.482C>T	p.Ala161Val	EGF4	Highly conserved (considering 26 species)	Affect protein function (score 0.01)	Probably Damaging (score 0.995)	2	RP with PPRPE		(den Hollander, et al., 1999)
2	c.584G>T	p.Cys195Phe	EGF5	Highly conserved (considering 26 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.998)	1	RP with PPRPE		(den Hollander, et al., 2004)
2	c.613_619del	p.Ile205Aspfs*13	EGF5	protein	-	-	14	LCA		(den Hollander, et al.,

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
				truncation, NMD				EORD		2001a; Galvin, et al., 2005; Hanein, et al., 2004; Lotery, et al., 2001a; Vallespin, et al., 2007b; Zernant, et al., 2005) this study (CIC00229)
3	c.717_718insG	Gln240Alafs*21	EGF6	protein truncation, NMD	-	-	1	LCA		(Henderson, et al., 2010)
3	c.750T>G	p.Cys250Trp	EGF6	Highly conserved (considering 24 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.918)	6	LCA, EORCD, EOCRD, PPRPE, ret talangiectasia		(den Hollander, et al., 1999; Henderson, et al., 2010; Henderson, et al., 2007)
4	c.915T>A	p.Cys305*	EGF8	protein truncation, NMD	-	-	1	RP	no cosegregation or phenotype information	(Vallespin, et al., 2007a)
4	c.929G>A	p.Cys310Tyr	EGF8	Highly conserved (considering 22 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.940)	1	EORD		(Coppieters, et al., 2010)
4	c.936T>G	p.Asn312Lys	EGF8	Moderately conserved (considering 22 species, His in Squirrel, Hedgehog, Tetraodon)	Affect protein function (score 0.01)	Benign (score 0.071)	1	EOCRD, ret talangiectasia		(Henderson, et al., 2010)
5	c.998G>A	p.Gly333Asp	EGF8	Highly conserved (considering 21 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.997)	2	LCA		(Seong, et al., 2008)
5	c.1084C>T	p.Gln362*	EGF9	protein truncation, NMD	-	-	5	LCA, EORD		(Coppieters, et al., 2010; den Hollander, et al., 2007; Yzer, et al., 2006)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
5	c.1125C>G	p.Tyr375*	EGF9	protein truncation, NMD	-	-	2	EORD, nanophthalmos		(Zenteno, et al., 2011)
5	c.1148G>A	p.Cys383Tyr	EGF9	Highly conserved (considering 22 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.999)	1	LCA		(Lotery, et al., 2001a)
6	c.1208C>G	p.Ser403*	EGF10	protein truncation, NMD	-	-	2	RP PPRPE, RP, ret talangiectasia		(den Hollander, et al., 2001b; den Hollander, et al., 1999)
6	**het.c.1269C>A,**	**p.Cys423***	**EGF10**	**protein truncation, NMD**	**-**	**-**	**1**	**EORD**	**(not found in 362 control alleles)**	**This study**
6	c.1298A>G	p.Tyr433Cys (!)	EGF10	Moderately conserved (considering 24 species, Phe in Cow, Elephant)	Affect protein function (score 0.04)	Probably Damaging (score 0.881)	1	RP, ret talangiectasia	(!) A stop mutation was present on the same allele (p.Ser403*)	(den Hollander, et al., 2001b)
6	c.1313G>A	p.Cys438Tyr	EGF10	Highly conserved (considering 23 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.998)	1	LCA PPRPE		(Simonelli, et al., 2007)
6	c.1438T>C	p.Cys480Arg	EGF11	Highly conserved (considering 25 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.998)	2	LCA		(Galvin, et al., 2005; Lotery, et al., 2001b)
6	c.1438T>G	p.Cys480Gly	EGF11	Highly conserved (considering 25 species)	Affect protein function (score 0.01)	Probably Damaging (score 0.997)	2	LCA		(Lotery, et al., 2001b)
6	c.1576C>T	p.Arg526*	LamAG 1	protein truncation, NMD	-	-	2	LCA		(Henderson, et al., 2010; Seong, et al., 2008)
6	c.1604T>C	p.Leu535Pro	LamAG 1	Moderately conserved (considering 26	Tolerated (score 0.08)	Probably Damaging (score	1	LCA		(Vallespin, et al., 2007b)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
				species; Met in Squirrel)		0.999)				
6	c.1690G>T	p.Asp564Tyr	LamAG 1	Highly conserved (considering 23 species)	Affect protein function (score 0.02)	Probably Damaging (score 0.998)	1	LCA		(Vallespin, et al., 2007b)
6	c.1733T>A	p.Val578Glu	LamAG 1	Moderately conserved (considering 23 species, Leu in Mouse and X. tropicalis)	Tolerated (score 0.27)	Probably Damaging (score 0.852)	1	RP, ret talangiectasia		(den Hollander, et al., 2004)
6	c.1750G>T	p.Asp584Tyr	LamAG 1	Weakly conserved (considering 23 species)	Tolerated (score 0.15)	Probably Damaging (score 0.941)	3	LCA, EORD	considered as likely pathogenic due to cosegregation in the family	(Hanein, et al., 2004) This study
6	c.1760G>A	p.Cys587Tyr	LamAG 1	Highly conserved up to Lizard (considering 20 species)	Affect protein function (score 0.04)	Probably Damaging (score 0.999)	1	RP, ret talangiectasia		(den Hollander, et al., 2004)
6	c.1834T>C	p.Ser611Pro	LamAG 1	Highly conserved in primates		Possibly Damaging (score 0.765)	4	LCA		(Li, et al., 2011)
6	**c.1963delC**	**p.Leu655Trpfs*10**	**LamAG 1**	**protein truncation, NMD**	**-**	**-**	**1**	**EORD**	**unknown second allele (not found in 376 control alleles)**	**This study**
6	c.2025G>T	p.Trp675Cys	EGF12	Moderately conserved up to Lizard (considering 20 species, Pro in Mouse lemur)	Tolerated (score 0.16)	Probably Damaging (score 0.997)	1	RP, ret talangiectasia		(Henderson, et al., 2010)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
6	c.2042G>A	p.Cys681Tyr	EGF12	Highly conserved (considering 24 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.999)	3		In Henderson at al this mutation was denoted as c.2043G>A, p.Cys681*	(Galvin, et al., 2005; Henderson, et al., 2010; Lotery, et al., 2001a)
6	c.2128G>C	p.Glu710Gln	LamininAG 2(den Hollander, et al., 2004)	Highly conserved (considering 22 species)	Tolerated (score 0.44)	Possibly Damaging (score 0.736)	3	LCA		(Hanein, et al., 2004)
7	c.2129C>T	p.Glu710Val	LamininAG 2(den Hollander, et al., 2004)	Highly conserved (considering 22 species)	Tolerated (score 0.20)	Probably Damaging (score 0.869)	4	RP		(Clark, et al., 2010; Henderson, et al., 2010)
7	c.2185_2186ins Alu	codon729 insAlu	LamAG 2	frameshift, NMD	-	-	2	RP PPRPE		(den Hollander, et al., 1999)
7	**c.2219C>T**	**p.Ser740Phe**	**amAG 2**	**Highly conserved (considering 26 species)**		**Probably Damaging (score 0.981)**	**2**	**RP**	**consanguinous family, detected by NGS (not found in 362 control alleles)**	**This study**
7	c.2222T>C	p.Met741Thr	LamAG 2	Highly conserved up to Lizard (considering 21 species)	Tolerated (score 0.19)	Possibly Damaging (score 0.832)	4	LCA, EORD		(Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Li, et al., 2011) This study
7	c.2234C>T	p.Thr745Met	LamAG 2	Highly conserved (considering 25 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.996)	15	LCA, RP, PPRPE, EORCD ret talangiectasia		(Clark, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 1999; Hanein, et al., 2004; Henderson, et al., 2010; Simonelli, et al., 2007; Yzer, et al., 2006)
7	c.2245_2247del 3bp (TCA)	p.Ser749del	LamAG 2	Weakly conserved (considering 25 species)	-	-	5	LCA, EORD, RP PPRPE		(Bernal, et al., 2003; Jacobson, et al., 2003; Tosi, et al., 2009; Vallespin, et al., 2007b)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
7	c.2258T>C	p.Leu753Pro	LamAG 2	Highly conserved up to Chicken (considering 20 species)	Tolerated (score 0.14)	Probably Damaging (score 0.994)	1	LCA, ret talangiectasia	p.Phe488Ser mutation on the second allele, which didn't co-segregate in the family	(Galvin, et al., 2005)
7	c.2290C>T	p.Arg764Cys	LamAG 2	Not conserved (considering 25 species)	Tolerated (score 0.23)	Benign (score 0.015)	16	LCA, EORD, RP PPRPE, ret talangiectasia	This change has been considered as likely pathogenic regardless poor conservation and low pathogenicity predictions. The decision was based on the genetic data - cosegregation, lack in the control alleles	(Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001b; den Hollander, et al., 1999; Galvin, et al., 2005; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Lotery, et al., 2001a; Vallespin, et al., 2007b)
7	**c.2365_2367delAAT, in frame deletion**	**p.Asn789del**	**LamAG 2**	**Not conserved (considering 24 species, Ser in Tarsier)**	**-**	**-**	**1**	**EORD**	**this inframe deletion is likely pathogenic, because it co-segregates in the family (not found in 362 alleles)**	**This study**
7	c.2401A>T	p.Lys801*	LamAG 2	protein truncation, NMD	-	-	27	LCA, RP, EORD, PPRPE, ret talangiectasia		(Booij, et al., 2005; Clark, et al., 2010; Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001a; den Hollander, et al., 2001b; Galvin, et al.,

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
										2005; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Simonelli, et al., 2007; Yzer, et al., 2006)
7	c.2438_2439ins >100A	insertion of >100 bp poly A, codons 812-813	LamAG 2	frameshift, NMD	-	-	1	LCA	unknown second allele	(Lotery, et al., 2001a)
7	c.2441_2442del	p.Leu814Argfs*23	LamAG 2	protein truncation, NMD	-	-	1	LCA		(Coppieters, et al., 2010)
7	c.2465G>A	p.Trp822*	LamAG 2	protein truncation, NMD	-	-	2	EORP, EORP PPRPE		(Riveiro-Alvarez, et al., 2008; Vallespin, et al., 2007b)
7	c.2479G>T	p.Gly827*	LamAG 2	protein truncation, NMD	-	-	1	LCA		(Hanein, et al., 2004)
7	c.2506C>A	p.Pro836Thr	LamAG 2	Highly conserved up to chicken (considering 17 species)	Tolerated (score 0.60)	Probably Damaging (score 0.991)	6	EORD, EOCRD, RP PPRPE		(den Hollander, et al., 2004; Henderson, et al., 2010) This study
7	c.2509G>C	p.Asp837His*	LamAG 2	Weakly conserved (considering 22 species)	Tolerated (score 0.28)	Possibly Damaging (score 0.604)	1	RP ret telangiectasia	two mutations on the same allele (with p.Ala1354Thr), cosegregation	(den Hollander, et al., 2001a)
7	c.2536G>A	p.Gly846Arg	LamAG 2	Highly conserved (considering 22 species)	Tolerated (score 0.35)	Probably Damaging (score 0.997)	4	EORP, RP PPRPRE		(Henderson, et al., 2010; Khaliq, et al., 2003)
7	c.2548_2551del GGCT	p.Gly850Valfs*5	LamAG 2	protein truncation, NMD	-	-	2	LCA	unknown second allele	(Galvin, et al., 2005; Lotery, et al., 2001a)
7	c.2548G>A	p.Gly850Ser	LamAG 2	Highly conserved (considering 22 species)	Tolerated (score 0.09)	Probably Damaging (score 0.995)	6	LCA, RP, RP PPRPE		(Clark, et al., 2010; den Hollander, et al., 2004; Henderson, et al., 2010)
7	c.2555T>C	p.Ile852Thr	LamAG 2	Weakly conserved	Tolerated (score 0.23)	Possibly Damaging	2	LCA, RP		(Hanein, et al., 2004; Simonelli, et al., 2007)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
				(considering 22 species, Val in Bushbaby, Mouse, Horse)		(score 0.426)				
7	c.2611_2613insT	p. Asn871Ilefs*38 or p.Ala872Cysfs*37	LamAG 2	protein truncation, NMD	-	-	1	LCA	originally it was reported as insT in 871 codon	(Lotery, et al., 2001a)
7	c.2671T>G	p.Cys891Gly	EGF13	Highly conserved (considering 21 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.993)	1	EORP		(Bernal, et al., 2003)
7	c.2676delG	p.Lys892Asnfs*16	EGF13	protein truncation, NMD	-	-	2	LCA	Originally reported as p.Lys892Asnfs*95	(Henderson, et al., 2010)
8	c.2681A>G	p.Asn894Ser	EGF13	Weakly conserved (considering 24 species, Ser in Platypus)	Tolerated (score 0.56)	Benign (score 0.017)	2	EORP, RP ret telangiectasia	unknown second allele in both cases, co-segregates in two affected family members (den Hollander, et al., 2001a)	(den Hollander, et al., 2001a; Vallespin, et al., 2007b)
8	c.2688T>A	p.Cys896*	EGF13	protein truncation, NMD	-	-	9	LCA, RP, EORP ret telangiectasia		(Hanein, et al., 2004; Henderson, et al., 2010; Vallespin, et al., 2007b; Yzer, et al., 2006)
8	c.2816G>A	p.Cys939Tyr	EGF14	Highly conserved (considering 24 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.995)	2	LCA		(den Hollander, et al., 2007)
9	c.2843G>A	p.Cys948Tyr	EGF14	Highly conserved (considering 22 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.995)	96	LCA, EORD, EOCRD, ret telangiectasia, PPRPE		(Bernal, et al., 2003; Booij, et al., 2005; Clark, et al., 2010; Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001a;

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
										den Hollander, et al., 2007; den Hollander, et al., 1999; Galvin, et al., 2005; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Lotery, et al., 2001a; Riveiro-Alvarez, et al., 2008; Tosi, et al., 2009; Vallespin, et al., 2007b; Yzer, et al., 2006; Zernant, et al., 2005) This study
9	c.2853dupT	p.Ala952Cysfs*4	EGF14 or LamininAG 3	protein truncation, NMD	-	-	2	LCA		(Hanein, et al., 2004)
9	c.2884_2886 delTTA	p.Leu962del	LamAG 3	Weakly conserved (considering 23 species)	-	-	1	EORP, choroideremia like fundus	unknown second allele	(Bernal, et al., 2003)
9	c.2957A>T	p.Asn986Ile	LamAG 3	Weakly conserved (considering 25 species)	Tolerated (score 0.17)	Possibly Damaging (score 0.744)	1	RP PPRPE	considered as likely pathogenic due to cosegregation in the family	(den Hollander, et al., 2004)
9	c.2966T>C	p.Ile989Thr	LamAG 3	Highly conserved in placental mammals (considering 17 species)	Tolerated (score 0.08)	Possibly Damaging (score 0.618)	2	LCA		(Khaliq, et al., 2003)
9	c.2983G>T	p.Glu995*	LamAG 3	protein truncation, NMD	-	-	1	LCA		(den Hollander, et al., 1999)
9	c.3002A>T	p.Ile1001Asn	LamAG 3	Moderately conserved up to Lizard (considering 26	Tolerated (score 0.37)	Probably Damaging (score 0.910)	2	LCA	considered as likely pathogenic due to cosegregation	(Vallespin, et al., 2007b)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
				species)					in the family	
9	c.3008T>C	p.Ile1003Thr	LamAG 3	Moderately conserved up to Lizard (considering 26 species)	Tolerated (score 0.08)	Probably Damaging (score 0.980)	1	LCA		(Henderson, et al., 2010)
9	c.3035T>C	p.Leu1012Ser	LamAG 3	Highly conserved (considering 26 species)	Tolerated (score 0.38)	Probably Damaging (score 0.995)	1	RP		(Henderson, et al., 2010)
9	c.3037C>T	p.Gln1013*	LamAG 3	protein truncation, NMD	-	-	1	EORD	unknown second allele	(Henderson, et al., 2010)
9	c.3074G>A	p.Ser1025Asn	LamAG 3	Moderately conserved (considering 25 species)	Tolerated (score 0.52)	Possibly Damaging (score 0.707)	2	RP ret telangiectasia	Originally reported as p.Ser1025Ala	(Henderson, et al., 2010)
9	c.3074G>T	p.Ser1025Ile	LamAG 3	Moderately conserved (considering 25 species)	Tolerated (score 0.19)	Probably Damaging (score 0.915)	2	LCA		(Hanein, et al., 2004)
9	c.3122T>C	p.Met1041Thr	LamAG 3	Highly conserved (considering 25 species)	Tolerated (score 0.40)	Probably Damaging (score 0.980)	2	RP PPRPE		(den Hollander, et al., 1999)
9	c.3212T>C	p.Leu1071Pro	LamAG 3	Highly conserved (considering 25 species)	Tolerated (score 0.23)	Probably Damaging (score 0.999)	4	RP PPRPE		(den Hollander, et al., 1999; Khaliq, et al., 2003)
9	c.3296C>A	p.Thr1099Lys	LamAG 3	Highly conserved up to Sloth (considering 17 species)	Tolerated (score 0.31)		2	RP		(Azam, et al., 2011)
9	c.3299T>C	p.Ile1100Thr	LamAG 3	Highly conserved up to Lizard (considering 21 species)	Tolerated (score 0.88)	Possibly Damaging (score 0.537)	8	LCA, EORP, RP PPRPE		(Vallespin, et al., 2007b)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
9	c.3299T>G	p.Ile1100Arg	LamAG 3	Highly conserved up to Lizard (considering 21 species)	Tolerated (score 0.53)	Probably Damaging (score 0.941)	1	LCA		(den Hollander, et al., 2001a)
9	c.3307G>A/C	p.Gly1103Arg	LamAG 3	Not conserved (considering 25 species)	Affect protein function (score 0.04)	Probably Damaging (score 0.852)	6	LCA, EORD		(Benayoun, et al., 2009; Hanein, et al., 2004; Simonelli, et al., 2007) This study
9	c.3320T>C	p.Leu1107Pro	LamAG 3	Highly conserved (considering 25 species)	Tolerated (score 0.24)	Probably Damaging (score 0.997)	2	LCA		(Hanein, et al., 2004; Henderson, et al., 2010)
9	c.3320T>G	p.Leu1107Arg	LamAG 3	Highly conserved (considering 25 species)	Tolerated (score 0.35)	Probably Damaging (score 0.997)	5	LCA		(Hanein, et al., 2004)
9	c.3331G>T	p.Glu1111*	LamAG 3	protein truncation, NMD	-	-	1	LCA		(den Hollander, et al., 2001a)
9	c.3343_3352del	p.Gly1115Ilefs*23	LamAG 3	protein truncation, NMD	-	-	2	EORP		(Lotery, et al., 2001a)
9	c.3347delT	p.Phe1116Serfs*25	LamAG 3	protein truncation, NMD	-	-	1	LCA		(Hanein, et al., 2004)
9	c.3427delT	p.Cys1143Alafs*67	EGF15	protein truncation, NMD	-	-	1	RP PPRPE		(den Hollander, et al., 2004)
9	c.3482A>G	p.Tyr1161Cys	EGF15	Moderately conserved (considering 25 species, His in Cow)	Affect protein function (score 0.01)	Probably Damaging (score 0.941)	1	No phenotype information	unknown second allele, no cosegregation information	(Vallespin, et al., 2010)
9	c.3493T>C	p.Cys1165Arg	EGF15	Highly conserved (considering 26 species)		Probably Damaging (score 0.999)	1	LCA		(Li, et al., 2011)
9	c.3655T>G	p.Cys1174Gly	EGF15	Highly conserved (considering 25 species)	Affect protein function	Probably Damaging (score	2	LCA, RP ret telangiectas		(Henderson, et al., 2010)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
					(score 0.00)	0.997)		ia		
9	c.3541T>C	p.Cys1181Arg	EGF16	Moderately conserved (considering 25 species, Tyr in Hedgehog)	Affect protein function (score 0.00)	Probably Damaging (score 0.999)	1	RP ret telangiectasia		(den Hollander, et al., 2001a)
9	c.3542dupG	p.Cys1181Trpfs*13	EGF16	frameshift, NMD	-	-	4	LCA/EORD		(Henderson, et al., 2010)
9	**c.3593A>G**	**p.Tyr1198Cys**	**EGF16**	**Moderately conserved (considering 25 species, Phe in Sloth and Tetraodon)**	**Affect protein function (score 0.02)**	**Probably Damaging (score 0.999)**	**1**	**EORD**	**not found in 378 control alleles**	**This study**
9	c.3613G>A	p.Gly1205Arg	EGF16	Highly conserved (considering 25 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.999)	1	LCA	unknown second allele	(Lotery, et al., 2001a)
9	c.3653G>T	p.Cys1218Phe	EGF17	Highly conserved (considering 25 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.998)	1	LCA/EORD		(Jacobson, et al., 2003)
9	**c.3659_3660delinsA**	**p.Ser1220Asnfs*62**	**EGF17**	**protein truncation, NMD**	**-**	**-**	**2**	**EORD**	**not found in 378 control alleles**	**This study**
9	c.3664C>T	p.Gln1222*	EGF17	protein truncation, NMD	-	-	1	LCA		(Yzer, et al., 2006)
9	**c.3668G>C**	**p.Cys1223Ser**	**EGF17**	**Highly conserved (considering 25 species)**	**Affect protein function (score 0.00)**	**Probably Damaging (score 0.995)**	**1**	**EORD**	**not found in 378 control alleles**	**This study**
9	c.3676G>T	p.Gly1226*	EGF17	protein truncation, NMD	-	-	3	LCA		(Li, et al., 2011)
9	c.3713_3716dup	p.Cys1240Profs*24	EGF17	protein truncation, NMD	-	-	1	LCA		(Coppieters, et al., 2010)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
11	c.3879G>A	p.Trp1293*	EGF18	protein truncation, NMD	-	-	4	LCA		(Coppieters, et al., 2010; Hanein, et al., 2004)
11	c.3914C>T	p.Pro1305Leu	EGF19	Moderately conserved (considering 25 species, Leu in Hedgehog)	Affect protein function (score 0.02)	Probably Damaging (score 1.00)	2	RP		(Siemiatkowska, et al., 2011)
11	c.3949A>C	p.Asn1317His	EGF19	Moderately conserved (considering 24 species)	Affect protein function (score 0.05)	Possibly Damaging (score 0.840)	1	LCA	unknown second allele	(Lotery, et al., 2001a)
11	c.3961T>A	p.Cys1321Ser	EGF19	Highly conserved (considering 24 species)	Affect protein function (score 0.00)	Possibly Damaging (score 0.849)	3	LCA, EORD		(Hanein, et al., 2004; Lotery, et al., 2001a)
11	c.3988delG	p.Glu1330Serfs*11	EGF19	protein truncation, NMD	-	-	1	LCA		(Hanein, et al., 2004)
11	c.3988G>T	p.Glu1330*	EGF19	protein truncation, NMD	-	-	2	LCA, ret telangiectasia		(Coppieters, et al., 2010; Vallespin, et al., 2007b)
11	c.3995G>T	p.Cys1332Phe	EGF19	Highly conserved (considering 24 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.998)	2	LCA		(den Hollander, et al., 2007)
11	c.3996C>A	p.Cys1332*	EGF19	protein truncation, NMD	-	-	1	LCA	unknown second allele	(Lotery, et al., 2001a)
11	c.3997G>T	p.Glu1333*	EGF19	protein truncation, NMD	-	-	1	LCA		(den Hollander, et al., 2001a)
12	c.4094C>A	p.Ala1365Asp	TM	Weakly conserved (considering 24 species)	Tolerated (score 0.10)	Possibly Damaging (score 0.762)	1	EORD	This variant was considered as likely pathogenic because of the change of the non-polar Ala in the hydrophobic	(Henderson, et al., 2010)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
									stretch to a polar Asp	
12	c.4121_4130del	p.Ala1374Glufs*20	C	protein truncation, NMD	-	-	5	LCA, EORD		(Benayoun, et al., 2009; Gerber, et al., 2002; Hanein, et al., 2004)
12	c.4142C>T	p.Pro1381Leu	C	Highly conserved (considering 25 species)	Affect protein function (score 0.00)	Probably Damaging (score 0.989)	1	LCA		(Henderson, et al., 2010)
12	c.4148G>A	p.Arg1383His	C	Moderately conserved (considering 25 species, Gly in Mouse, Trp in Hedgehog)	Tolerated (score 0.14)	Possibly Damaging (score 0.802)	2	RP, RP with PPRPE	unknown second allele	(Clark, et al., 2010; den Hollander, et al., 2004)
IVS6	c.2128+2T>G	-	-	splicing alteration, NMD	-	-	1			(Li, et al., 2011)
IVS8	c.2842+5G>A	-	-	splicing alteration, NMD	-	-	9	LCA, RP, PPRPE, Ret telangiectasia		(Coppieters, et al., 2010; den Hollander, et al., 2001b; den Hollander, et al., 1999; Yzer, et al., 2006)
IVS10	c.3878+1G>T	-	-	splicing alteration, NMD	-	-	1	LCA		(den Hollander, et al., 2001a)
IVS11	c.4005+1G>A	-	-	splicing alteration, NMD	-	-	3	LCA		(Coppieters, et al., 2010; Hanein, et al., 2004)
IVS11	c.4005+2T>G	-	-	splicing alteration, NMD	-	-	4	LCA		(Li, et al., 2011)
IVS11	c.4006-2A>G	-	-	splicing alteration, NMD	-	-	1	LCA		(Li, et al., 2011)
IVS11	c.4006-1G>T	-	-	splicing alteration, NMD	-	-	1	LCA		(Coppieters, et al., 2010)
	no second allele	no second allele	no second allele	no second allele			70	LCA, RP, PPRPE, ret telangiectasia		(Bernal, et al., 2003; Booij, et al., 2005; Clark, et al., 2010; den Hollander, et al., 2004;

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen predictions	No. of reported alleles	Phenotype	remarks	reference
										den Hollander, et al., 2001b; Galvin, et al., 2005; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Li, et al., 2011; Lotery, et al., 2001a; Simonelli, et al., 2007; Vallespin, et al., 2010; Vallespin, et al., 2007b; Zernant, et al., 2005)

Novel mutations are presented in bold. Nucleotide numbering is based on cDNA sequence from the Ref. NM_201253.2, where A of the ATG initiation codon is 1. Lam AG – Laminin AG like domain; TM – transmembrane; C – cytoplasmic; LCA –Leber congenital amaurosis; RP – retinitis pigmentosa; EORD – early onset retinal dystrophy; PPRPE – preservation of para-arteriolar retinal pigment epithelium. For PolyPhen-2 the HumVar value was taken, which is preferred for the diagnostic of human Mendelian diseases. In the conservation analysis the following species were considered: Human, Chimp, Gorilla, Rhesus, Tarsier, Mouse lemur, Bushbaby, Tree shrew, Mouse, Squirrel, Rabbit, Cow, Horse, Cat, Dog, Hedgehog, Elephant, Sloth, Wallaby, Opossum, Platypus, Chicken, Lizard, X. tropicalis, Tetraodon, Stickleback and Zebrafish. The conservation criteria have been described in the Supp. Methods.

Supp. Table S2. Unclassified nonsynonymous changes

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen -2 predictions	No. of reported alleles	Phenotype	remarks	reference
2	c.619G>A	p.Val162Met	EGF4	Weakly conserved (considering 27 species; Met in Mouse, Rabbit, Cow, Dog)	Tolerated (score 0.25)	Benign (score 0.023)	1	PPCRA	Dominant inheritance, not present in 150 controls, co-segregates in the family, LOD score: 1.8	(McKay, et al., 2005)
2	c.614T>C	p.Ile205Thr	EGF5	Moderately conserved in vertebrates considering 26 species (Val in Mouse Lemur, Opossum and Stickleback)	Tolerated (score 0.45)	Possibly Damaging (score 0.629)	5	LCA, EORD, RP	For this variant the second mutant *CRB1* allele has never been shown. It has been suggested as non-pathogenic (den Holl2004). Cosegregation of this change has been shown with a mutant allele from another parent (Vallespin 2007). Digenic inheritence with *GUCY2D* and *RPGRIP1* have been suggested in Villespin et al but the digenic mutations did not co-segregate.	(Bernal, et al., 2003; den Hollander, et al., 2004; Henderson, et al., 2010; Vallespin, et al., 2007b)
6	c.1472A>T	p.Asp491Val	LamA G 1	Not conserved	Tolerated (score 0.28)	Benign (score 0.090)	1	EORD	Considered as unclassified variant by the authors, p.Cys948Tyr was present on the second allele, no cosegregation information	(Coppieters, et al., 2010)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen -2 predictions	No. of reported alleles	Phenotype	remarks	reference
6	c.1903T>C	p.Ser635Pro	LamAG 1	Weakly conserved (considering 27 species; Pro in Mouse lemour)	Tolerated (score 0.17)	Benign (score 0.047)	1	LCA	Second mutation is a likeley pathogenic splice mutation, however no cosegregation analysis was performed	(Li, et al., 2011)
8	c.2809G>A	p.Ala937Thr	EGF14	Highly conserved in placental mammals (considering 16 species)	Tolerated (score 0.13)	Possibly Damaging (score 0.838)	1	LCA	Considered as polymorphism by the authors, however it was not present in 170 controls, no cosegregation data was available. Due to high conservation and Polyphen2 prediction it is considered as unclassified variant	(Seong, et al., 2008)
9	c.3103C>T	p.His1035Tyr	LamAG 3	Moderately conserved (considering 25 species, Tyr in Cow)	Tolerated (score 1.00)	Benign (score 0.027)	1	LCA/RP?	Unknown second allele, not found in 100 controls, no cosegregation information	(Henderson, et al., 2010)
11 alt	c.4082G>A	p.Arg1361His	TM	Moderately conserved (in this case, conservation of the Arg codon (CGT) was considered in 23 species; in Hedgehog and Stickleback the CAC codes for His)	this alternative transcript failed to be analysed	Benign (score 0.010)	1	LCA	Unknown second allele; mutation in the alternative transcript AF154671	(Simonelli, et al., 2007)

Exon	Nucleotide change	Aminoacid change	Protein domain	Effect/residue conservation	SIFT predictions	PolyPhen -2 predictions	No. of reported alleles	Phenotype	remarks	reference
12	c.4060G>A	p.Ala1354Thr	TM	Moderately conserved up to X. tropicalis (considering 22 species, Val in Mouse lemour and dog)	Tolerated (score 0.11)	Benign (score 0.180)	1	RP ret telangiectasia	Second mutation on the same allele (p.Asp837His)	(den Hollander, et al., 2001a)

Nucleotide numbering is based on cDNA sequence from the Ref. NM_201253.2, where A of the ATG initiation codon is 1. Lam AG – Laminin AG like domain; TM – transmembrane; C – cytoplasmic. LCA –Leber congenital amaurosis; RP – retinitis pigmentosa; EORD – early onset retinal dystrophy; PPRPE – preservation of para-arteriolar retinal pigment epithelium. For PolyPhen-2 the HumVar value was taken, which is preferred for the diagnostic of human Mendelian diseases. In the conservation analysis the following species were considered: Human, Chimp, Gorilla, Rhesus, Tarsier, Mouse lemur, Bushbaby, Tree shrew, Mouse, Squirrel, Rabbit, Cow, Horse, Cat, Dog, Hedgehog, Elephant, Sloth, Wallaby, Opossum, Platypus, Chicken, Lizard, X. tropicalis, Tetraodon, Stickleback and Zebrafish. The conservation criteria have been described in the Supp. Methods.

Supp. Table S3. Unlikely pathogenic non-synonymous *CRB1* variants

Exon	nucleotide change	amino acid change	Protein domain	conservation	SIFT	PolyPhen	Comment	reference
4	c.866C>T	p.Thr289Met	EGF7	not conserved, Met in Elephant	Tolerated (score 0.18)	Benign (0.006)	no cosegregation	(Bernal, et al., 2003; den Hollander, et al., 2001a; Lotery, et al., 2001a; Simonelli, et al., 2007; Vallespin, et al., 2007b)
6	c.1463T>C	p.Phe488Ser	LamAG-1	conserved	Tolerated (score 0.09)	Probably Damaging (score 0.992)	reported as a second mutant allele to the p.Leu753Pro mutation, but p.Phe488Ser did not co-segregate in the family	(Galvin, et al., 2005)
6	c.2035C>G	p.Gln679Glu	EGF12	not conserved	Tolerated (score 0.25)	Possibly Damaging (score 0.616)	no cosegregation	(Bernal, et al., 2003; den Hollander, et al., 2004)
7	c.2306_2307GC>AG	p.Arg769Gln	LamAG-2	not conserved	Tolerated (score 0.22)	Benign (0.003)	present in control alleles, no cosegregation and no second *CRB1* mutation found	(Bernal, et al., 2003; Lotery, et al., 2001a; Vallespin, et al., 2007b; Zernant, et al., 2005)
7	c.2306G>A	p.Arg769His	LamAG-2	not conserved, His in Rhesus	Tolerated (score 0.39)	Benign (0.001)	-	(Bernal, et al., 2003; Seong, et al., 2008)

Exon	nucleotide change	amino acid change	Protein domain	conservation	SIFT	PolyPhen	Comment	reference
7	Not reported	p.Thr821Met	LamAG-2	not conserved	Tolerated (score 0.18)	Possibly Damaging (score 0.679)	no cosegregation	(den Hollander, et al., 2001a)
8	c.2714G>A	p.Arg905Gln	EGF13	not conserved	Tolerated (score 0.31)	Benign (0.063)	Digenism suspected with *RPGRIP1*, no cosegregation	(den Hollander, et al., 2004; Vallespin, et al., 2007b; Zernant, et al., 2005)
9	c.2875G>A	p.Gly959Ser	LamAG 3	not conserved, Ser in Rhesus	Tolerated (score 0.93)	Benign (score 0.000)	Only one reported allele, unknown second allele, no cosegregation information, not present in 372 controls. Originally it was classified as likely pathogenic	(den Hollander, et al., 2004)
11	c.3992G>A	p.Arg1331His	EGF19	Highly conserved up to Opossum (considering 16 species, His in Platypus, X.tropicalis, Stickleback)	Tolerated (score 0.69)	Benign (0.131)	Present in control alleles, no cosegregation and no second *CRB1* mutation ever documented	den (Bernal, et al., 2003; den Hollander, et al., 2001a; Lotery, et al., 2001a; Vallespin, et al., 2007b)

Nucleotide numbering is based on cDNA sequence from the Ref. NM_201253.2, where A of the ATG initiation codon is 1. Lam AG – Laminin AG like domain; TM – transmembrane; C – cytoplasmic. For PolyPhen-2 the HumVar value was taken, which is preferred for the diagnostic of human Mendelian diseases. In the conservation analysis the following species were considered: Human, Chimp, Gorilla, Rhesus, Tarsier, Mouse lemur, Bushbaby, Tree shrew, Mouse, Squirrel, Rabbit, Cow, Horse, Cat, Dog, Hedgehog, Elephant, Sloth, Wallaby, Opossum, Platypus, Chicken, Lizard, X. tropicalis, Tetraodon, Stickleback and Zebrafish. The conservation criteria have been described in the Supp. Methods.

Supp. References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nat Methods 7:248-9.

Audo I, Sahel JA, Mohand-Said S, Lancelot ME, Antonio A, Moskova-Doumanova V, Nandrot EF, Doumanov J, Barragan I, Antinolo G, Bhattacharya SS, Zeitz C. 2010. EYS is a major gene for rod-cone dystrophies in France. Hum Mutat 31:E1406-35.

Azam M, Collin RW, Malik A, Khan MI, Shah ST, Shah AA, Hussain A, Sadeque A, Arimadyo K, Ajmal M, Azam A, Qureshi N, Bokhari H, Strom TM, Cremers FP, Qamar R, den Hollander AI. 2011. Identification of novel mutations in pakistani families with autosomal recessive retinitis pigmentosa. Arch Ophthalmol 129:1377-8.

Benayoun L, Spiegel R, Auslender N, Abbasi AH, Rizel L, Hujeirat Y, Salama I, Garzozi HJ, Allon-Shalev S, Ben-Yosef T. 2009. Genetic heterogeneity in two consanguineous families segregating early onset retinal degeneration: the pitfalls of homozygosity mapping. Am J Med Genet A 149A:650-6.

Bernal S, Calaf M, Garcia-Hoyos M, Garcia-Sandoval B, Rosell J, Adan A, Ayuso C, Baiget M. 2003. Study of the involvement of the RGR, CRPB1, and CRB1 genes in the pathogenesis of autosomal recessive retinitis pigmentosa. J Med Genet 40:e89.

Booij JC, Florijn RJ, ten Brink JB, Loves W, Meire F, van Schooneveld MJ, de Jong PT, Bergen AA. 2005. Identification of mutations in the AIPL1, CRB1, GUCY2D, RPE65, and RPGRIP1 genes in patients with juvenile retinitis pigmentosa. J Med Genet 42:e67.

Clark GR, Crowe P, Muszynska D, O'Prey D, O'Neill J, Alexander S, Willoughby CE, McKay GJ, Silvestri G, Simpson DA. 2010. Development of a diagnostic genetic test for simplex and autosomal recessive retinitis pigmentosa. Ophthalmology 117:2169-77 e3.

Coppieters F, Casteels I, Meire F, De Jaegere S, Hooghe S, van Regemorter N, Van Esch H, Matuleviciene A, Nunes L, Meersschaut V, Walraedt S, Standaert L, Coucke P, Hoeben H, Kroes HY, Vande Walle J, de Ravel T, Leroy BP, De Baere E. 2010. Genetic screening of LCA in Belgium: predominance of CEP290 and identification of potential modifier alleles in AHI1 of CEP290-related phenotypes. Hum Mutat 31:E1709-66.

den Hollander AI, Davis J, van der Velde-Visser SD, Zonneveld MN, Pierrottet CO, Koenekoop RK, Kellner U, van den Born LI, Heckenlively JR, Hoyng CB, Handford PA, Roepman R, Cremers FP. 2004. CRB1 mutation spectrum in inherited retinal dystrophies. Hum Mutat 24:355-69.

den Hollander AI, Heckenlively JR, van den Born LI, de Kok YJ, van der Velde-Visser SD, Kellner U, Jurklies B, van Schooneveld MJ, Blankenagel A, Rohrschneider K, Wissinger B, Cruysberg JR, Deutman AF, Brunner HG, Apfelstedt-Sylla E, Hoyng CB, Cremers FP. 2001a. Leber congenital amaurosis and retinitis pigmentosa with Coats-like exudative

vasculopathy are associated with mutations in the crumbs homologue 1 (CRB1) gene. Am J Hum Genet 69:198-203.

den Hollander AI, Johnson K, de Kok YJ, Klebes A, Brunner HG, Knust E, Cremers FP. 2001b. CRB1 has a cytoplasmic domain that is functionally conserved between human and Drosophila. Hum Mol Genet 10:2767-73.

den Hollander AI, Lopez I, Yzer S, Zonneveld MN, Janssen IM, Strom TM, Hehir-Kwa JY, Veltman JA, Arends ML, Meitinger T, Musarella MA, van den Born LI, Fishman GA, Maumenee IH, Rohrschneider K, Cremers FP, Koenekoop RK. 2007. Identification of novel mutations in patients with Leber congenital amaurosis and juvenile RP by genome-wide homozygosity mapping with SNP microarrays. Invest Ophthalmol Vis Sci 48:5690-8.

den Hollander AI, ten Brink JB, de Kok YJ, van Soest S, van den Born LI, van Driel MA, van de Pol DJ, Payne AM, Bhattacharya SS, Kellner U, Hoyng CB, Westerveld A, Brunner HG, Bleeker-Wagemakers EM, Deutman AF, Heckenlively JR, Cremers FP, Bergen AA. 1999. Mutations in a human homologue of Drosophila crumbs cause retinitis pigmentosa (RP12). Nat Genet 23:217-21.

Galvin JA, Fishman GA, Stone EM, Koenekoop RK. 2005. Evaluation of genotype-phenotype associations in leber congenital amaurosis. Retina 25:919-29.

Gerber S, Perrault I, Hanein S, Shalev S, Zlotogora J, Barbet F, Ducroq D, Dufier J, Munnich A, Rozet J, Kaplan J. 2002. A novel mutation disrupting the cytoplasmic domain of CRB1 in a large consanguineous family of Palestinian origin affected with Leber congenital amaurosis. Ophthalmic Genet 23:225-35.

Hanein S, Perrault I, Gerber S, Tanguy G, Barbet F, Ducroq D, Calvas P, Dollfus H, Hamel C, Lopponen T, Munier F, Santos L, Shalev S, Zafeiriou D, Dufier JL, Munnich A, Rozet JM, Kaplan J. 2004. Leber congenital amaurosis: comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis. Hum Mutat 23:306-17.

Henderson RH, Mackay DS, Li Z, Moradi P, Sergouniotis P, Russell-Eggitt I, Thompson DA, Robson AG, Holder GE, Webster AR, Moore AT. 2010. Phenotypic variability in patients with retinal dystrophies due to mutations in CRB1. Br J Ophthalmol.

Henderson RH, Waseem N, Searle R, van der Spuy J, Russell-Eggitt I, Bhattacharya SS, Thompson DA, Holder GE, Cheetham ME, Webster AR, Moore AT. 2007. An assessment of the apex microarray technology in genotyping patients with Leber congenital amaurosis and early-onset severe retinal dystrophy. Invest Ophthalmol Vis Sci 48:5684-9.

Jacobson SG, Cideciyan AV, Aleman TS, Pianta MJ, Sumaroka A, Schwartz SB, Smilko EE, Milam AH, Sheffield VC, Stone EM. 2003. Crumbs homolog 1 (CRB1) mutations result in a thick human retina with abnormal lamination. Hum Mol Genet 12:1073-8.

Khaliq S, Abid A, Hameed A, Anwar K, Mohyuddin A, Azmat Z, Shami SA, Ismail M, Mehdi SQ. 2003. Mutation screening of Pakistani families with congenital eye disorders. Exp Eye Res 76:343-8.

Li L, Xiao X, Li S, Jia X, Wang P, Guo X, Jiao X, Zhang Q, Hejtmancik JF. 2011. Detection of variants in 15 genes in 87 unrelated chinese patients with leber congenital amaurosis. PLoS One 6:e19458.

Lotery AJ, Jacobson SG, Fishman GA, Weleber RG, Fulton AB, Namperumalsamy P, Heon E, Levin AV, Grover S, Rosenow JR, Kopp KK, Sheffield VC, Stone EM. 2001a. Mutations in the CRB1 gene cause Leber congenital amaurosis. Arch Ophthalmol 119:415-20.

Lotery AJ, Malik A, Shami SA, Sindhi M, Chohan B, Maqbool C, Moore PA, Denton MJ, Stone EM. 2001b. CRB1 mutations may result in retinitis pigmentosa without para-arteriolar RPE preservation. Ophthalmic Genet 22:163-9.

McKay GJ, Clarke S, Davis JA, Simpson DA, Silvestri G. 2005. Pigmented paravenous chorioretinal atrophy is associated with a mutation within the crumbs homolog 1 (CRB1) gene. Invest Ophthalmol Vis Sci 46:322-8.

McKibbin M, Ali M, Mohamed MD, Booth AP, Bishop F, Pal B, Springell K, Raashid Y, Jafri H, Inglehearn CF. 2010. Genotype-phenotype correlation for leber congenital amaurosis in Northern Pakistan. Arch Ophthalmol 128:107-13.

Ng PC, Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812-4.

Riveiro-Alvarez R, Vallespin E, Wilke R, Garcia-Sandoval B, Cantalapiedra D, Aguirre-Lamban J, Avila-Fernandez A, Gimenez A, Trujillo-Tiebas MJ, Ayuso C. 2008. Molecular analysis of ABCA4 and CRB1 genes in a Spanish family segregating both Stargardt disease and autosomal recessive retinitis pigmentosa. Mol Vis 14:262-7.

Seelow D, Schuelke M, Hildebrandt F, Nurnberg P. 2009. HomozygosityMapper--an interactive approach to homozygosity mapping. Nucleic Acids Res 37:W593-9.

Seong MW, Kim SY, Yu YS, Hwang JM, Kim JY, Park SS. 2008. Molecular characterization of Leber congenital amaurosis in Koreans. Mol Vis 14:1429-36.

Siemiatkowska AM, Arimadyo K, Moruz LM, Astuti GDN, Castro-Miro Md, Zonneveld MN, Strom TM, Wijs IJd, Hoefsloot LH, Faradz SMH, Cremers FPM, Hollander AId, Collin RWJ. 2011. Molecular genetic analysis of retinitis pigmentosa in Indonesia using genome-wide homozygosity mapping. Molecular Vision (in press).

Simonelli F, Ziviello C, Testa F, Rossi S, Fazzi E, Bianchi PE, Fossarello M, Signorini S, Bertone C, Galantuomo S, Brancati F, Valente EM, Ciccodicola A, Rinaldi E, Auricchio A, Banfi S. 2007. Clinical and molecular genetics of Leber's congenital amaurosis: a multicenter study of Italian patients. Invest Ophthalmol Vis Sci 48:4284-90.

Tosi J, Tsui I, Lima LH, Wang NK, Tsang SH. 2009. Case report: autofluorescence imaging and phenotypic variance in a sibling pair with early-onset retinal dystrophy due to defective CRB1 function. Curr Eye Res 34:395-400.

Vallespin E, Avila-Fernandez A, Velez-Monsalve C, Almoguera B, Martinez-Garcia M, Gomez-Dominguez B, Gonzalez-Roubaud C, Cantalapiedra D, Trujillo-Tiebas MJ, Ayuso C. 2010. Novel human pathological mutations. Gene symbol: CRB1. Disease: Leber congenital amaurosis. Hum Genet 127:119.

Vallespin E, Cantalapiedra D, Riveiro-Alvarez R, Aguirre-Lamban J, Avila-Fernandez A, Martinez MA, Gimenez A, Trujillo-Tiebas MJ, Ayuso C. 2007a. Human gene mutations. Gene symbol: CRB1. Disease: late onset retinitis pigmentosa. Hum Genet 122:212.

Vallespin E, Cantalapiedra D, Riveiro-Alvarez R, Wilke R, Aguirre-Lamban J, Avila-Fernandez A, Lopez-Martinez MA, Gimenez A, Trujillo-Tiebas MJ, Ramos C, Ayuso C. 2007b. Mutation screening of 299 Spanish families with retinal dystrophies by Leber congenital amaurosis genotyping microarray. Invest Ophthalmol Vis Sci 48:5653-61.

Yzer S, Leroy BP, De Baere E, de Ravel TJ, Zonneveld MN, Voesenek K, Kellner U, Ciriano JP, de Faber JT, Rohrschneider K, Roepman R, den Hollander AI, Cruysberg JR, Meire F, Casteels I, van Moll-Ramirez NG, Allikmets R, van den Born LI, Cremers FP. 2006. Microarray-based mutation detection and phenotypic characterization of patients with Leber congenital amaurosis. Invest Ophthalmol Vis Sci 47:1167-76.

Zenteno JC, Buentello-Volante B, Ayala-Ramirez R, Villanueva-Mendoza C. 2011. Homozygosity mapping identifies the Crumbs homologue 1 (Crb1) gene as responsible for a recessive syndrome of retinitis pigmentosa and nanophthalmos. Am J Med Genet A 155A:1001-6.

Zernant J, Kulm M, Dharmaraj S, den Hollander AI, Perrault I, Preising MN, Lorenz B, Kaplan J, Cremers FP, Maumenee I, Koenekoop RK, Allikmets R. 2005. Genotyping microarray (disease chip) for Leber congenital amaurosis: detection of modifier alleles. Invest Ophthalmol Vis Sci 46:3052-9.