CRB1 mutations in inherited retinal dystrophies.

Kinga Bujakowska, Isabelle Audo, Saddek Mohand-Saïd, Marie-Elise Lancelot, Aline Antonio, Aurore Germain, Thierry Léveillard, Mélanie Letexier, Jean-Paul Saraiva, Christine Lonjou, et al.

***CRB1* mutations in inherited retinal dystrophies**

Kinga Bujakowska[1,2,3], Isabelle Audo[1,2,3,4,5], Saddek Mohand-Saïd[1,2,3,4], Marie-Elise Lancelot[1,2,3], Aline Antonio[1,2,3,4], Aurore Germain[1,2,3], Thierry Léveillard[1,2,3], Mélanie Letexier[6], Jean-Paul Saraiva[6], Christine Lonjou[7], Wassila Carpentier[7], José-Alain Sahel[1,2,3,4,5,8], Shomi S. Bhattacharya[1,2,3,5,9], Christina Zeitz[1,2,3].

[1]INSERM, U968, Paris, F-75012, France; [2]CNRS, UMR_7210. Paris, F-75012, France; [3]UPMC Univ Paris 06, UMR_S 968, Department of Genetics, Institut de la Vision, Paris, F-75012, France; [4]Centre Hospitalier National d'Ophtalmologie des Quinze-Vingts, INSERM-DHOS CIC 503, Paris, F-75012, France; [5]UCL-Institute of Ophthalmology, London, UK; [6]IntegraGen SA, Genopole CAMPUS 1 bat G8 FR-91030 Evry, France; [7]Plateforme Post-génomique P3S, Hôpital Pitié Salpêtrière, Paris, France; [8]Fondation Ophtalmologique Adolphe de Rothschild, Paris, France; [9]Department of Cellular Therapy and Regenerative Medicine, Andalusian Centre for Molecular Biology and Regenerative Medicine (CABIMER), Isla Cartuja, Seville, Spain

Corresponding authors: Isabelle Audo and Christina Zeitz

Institut de la Vision

Department of Genetics

17, Rue Moreau, 75012 Paris

France

Email address: isabelle.audo@inserm.fr, christina.zeitz@inserm.fr

Financial Support: Foundation Fighting Blindness (I.A. FFB Grant No: CD-CL-0808-0466-CHNO and the CIC503 recognized as an FFB center, FFB Grant No: C-CMM-

0907-0428-INSERM04), Agence Nationale de la Recherche (SSB), Fondation Voir et Entendre (CZ), GIS-maladies rares (CZ), Ville de Paris and Région Ile de France, National Institutes of Health (USA) (KB NIH, Grant No: 1R01EY020902 - 01A1).

Abstract (max 200 words)

Mutations in the *CRB1* gene are associated with variable phenotypes of severe retinal dystrophies, ranging from Leber Congenital Amaurosis (LCA) to rod-cone dystrophy (also called retinitis pigmentosa (RP)). Moreover, retinal dystrophies resulting from *CRB1* mutations may be accompanied by specific fundus features: preservation of the para-arteriolar retinal pigment epithelium (PPRPE) and retinal telangiectasia with exudation (also referred to as Coats-like vasculopathy). In this publication we report seven novel mutations and classify over 150 reported *CRB1* sequence variants that were found in more that 240 patients. The data from previous reports was used to analyse a potential correlation between *CRB1* variants and the clinical features of respective patients. This meta-analysis suggests that the differential phenotype of patients with *CRB1* mutations is due to additional modifying factors rather than particular mutant allele combination.

Key words: CRB1, LCA, Retinitis Pigmentosa, rod-cone dystrophy

Background

Mutations in the *CRB1* gene (MIM#: 604210) are associated with variable phenotypes of severe retinal dystrophies, ranging from Leber Congenital Amaurosis (LCA) to rod-cone dystrophy (also called retinitis pigmentosa (RP)) (Benayoun, et al., 2009; 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; den Hollander, et al., 2007; den Hollander, et al., 1999; Galvin, et al., 2005; Gerber, et al., 2002; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Khaliq, et al., 2003; Li, et al., 2011; Lotery, et al., 2001a; Lotery, et al., 2001b; Riveiro-Alvarez, et al., 2008; Seong, et al., 2008; Simonelli, et al., 2007; Tosi, et al., 2009; Vallespin, et al., 2007; Walia, et al., 2010; Yzer, et al., 2006a; Yzer, et al., 2006b; Zernant, et al., 2005). LCA is a group of the most severe and the earliest occurring retinal dystrophies resulting in congenital blindness (den Hollander, et al., 2008). The onset of the disease occurs at birth and the characteristic features include non-recordable electroretinogram (ERG), nystagmus, sluggish or absent pupillary responses and oculo-digital reflexes, a distinctive eye-rubbing also called the Franschetti sign (den Hollander, et al., 2008; Franceschetti and Dieterle, 1954; Leber, 1869). RP is a clinically heterogeneous disorder characterised by a progressive degeneration of the photoreceptors and leading to a visual impairment of variable severity that can end in complete blindness. The disease onset is highly variable: it may commence in the first decade of life or much later. There is a considerable clinical overlap between LCA and early-onset RP and in some cases/reports the diagnosis is ambiguous. Early-onset RP, however, is considered as a relatively milder form, where patients do not have a congenital onset of visual impairment.

LCA and RP resulting from *CRB1* mutations may be accompanied by specific fundus features: preservation of the para-arteriolar retinal pigment epithelium (PPRPE) (Bernal, et al., 2003; den Hollander, et al., 2004; den Hollander, et al., 1999; Heckenlively, 1982; Henderson, et al., 2010; Khaliq, et al., 2003; Simonelli, et al., 2007; Yzer, et al., 2006b) and retinal telangiectasia with exudation (also referred to as Coats-like vasculopathy) (Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001a; Henderson, et al., 2010; Yzer, et al., 2006b). PPRPE is characterized by a relative preservation of retinal pigment epithelium (RPE) adjacent to retinal arterioles despite a panretinal RPE degeneration (Heckenlively, 1982). This is, however, not consistent in *CRB1*-associated RP and the absence of PPRPE in a severe RP should not exclude *CRB1* as a potential causal gene (Lotery, et al., 2001b). Retinal telangiectasia is a condition of abnormally permeable blood vessels, leading to exudation and retinal detachment (Cahill, et al., 2001). Some patients with *CRB1* mutations show macular atrophy (Henderson, et al., 2010), similar features were found for other LCA causing genes (*GUCY2D* MIM#:600179, *AIPL1* MIM#:604392 and *RPGRIP1* MIM#:605446), which lead to classification of LCA into cone-rod LCA and rod-cone LCA (Hanein, et al., 2004). Patients with *CRB1* mutations belong to both categories. Predisposition of the *CRB1* patients to keratoconus (McKibbin, et al., 2010; McMahon, et al., 2009) and implication for pigmented paravenous chorioretinal atrophy (McKay, et al., 2005) and nanophthalmos (Zenteno, et al., 2011) have also been reported.

CRB1 is a human homologue of the *Drosophila melanogaster* gene coding for protein crumbs (crb) and it is expressed in the retina and the brain (den Hollander, et al., 1999). *CRB1* consists of 12 exons and exhibits alternative splicing at the 3' end, yielding two proteins of 1376 and 1406 amino acids (den Hollander, et al., 2001b).

Both proteins contain 19 EGF-like domains, three laminin AG-like domains and a signal peptide sequence. In addition, the longer isoform contains transmembrane and cytoplasmic domains (den Hollander, et al., 2001b; Gosens, et al., 2008). The cytoplasmic domain includes conserved FERM and PDZ binding motifs, through which CRB1 participates in the formation of adherens junction and links to the actin cytoskeleton (Gosens, et al., 2008).

In *Drosophila*, crb determines the polarity of the embryonic epithelium and peripheral neurons; it is important for the maintenance of zonula adherens (ZA) and it is localized in the apical membrane (Tepass, et al., 1990). In the mouse retina, Crb1 is present in the apical membranes of the epithelial cells, in Muller cells and in photoreceptor inner segments, where it concentrates in the vicinity of the outer limiting membrane (den Hollander, et al., 2002; Mehalow, et al., 2003; Pellikka, et al., 2002; van de Pavert, et al., 2004). A similar distribution was found in the human retina (van de Pavert, et al., 2004). Crumbs and its mouse homolog Crb1 is involved in the photoreceptor morphogenesis (Pellikka, et al., 2002; Tepass, et al., 1990). Analysis of the naturally occurring *Crb1*^{rd8} mouse mutant, suggests a developmental defect of the retina, where disruption of the outer limiting membrane and formation of retinal folds (pseudorosettes) are observed (Mehalow, et al., 2003). Disorganization of the retinal layers was also noted in other *Crb1* mouse models (van de Pavert, et al., 2004; van de Pavert, et al., 2007). These findings are in accordance with clinical features of the patients carrying *CRB1* mutations, whose retinas are thickened and show an altered laminar organization, resembling an immature normal retina (Jacobson, et al., 2003). The latter further supports the importance of CRB1 in the development of the retina.

This study presents an overview of the previously published *CRB1* variants and novel mutations identified in a French cohort of simplex and autosomal recessive RP (arRP) patients. Based on the available genetic and phenotypic data from the literature and on our original findings, we classify all variants into one of the three groups (likely pathogenic, unclassified variants and unlikely pathogenic, Supp. Tables S1-S3). We discuss the clinical variability of patients harboring *CRB1* mutations and analyse the phenotype-genotype correlation of likely pathogenic changes. Identification of novel mutations in the French cohort is described (Supp. Methods and Results) and precise clinical characterisation is given.

Novel *CRB1* Variants

Eleven unrelated patients with ar or isolated RP in the French cohort carried likely pathogenic variants of *CRB1* (Table1). Seven mutations were novel: three missense changes (p.Ser740Phe, p.Tyr1198Cys and p.Cys1223Ser), one nonsense mutation (p.Cys423*), one in-frame deletion (p.Asn789del) and two frameshift deletions (p.Leu655Trpfs*10, p.Ser1220Asnfs*62) (Table 1). Mutations identified in this study were not present in the SNP databases nor listed as non-pathogenic variants in the literature. None of the novel mutations was present in at least 362 control alleles and the mutations co-segregated in available family members (Supp. Figure S1). In all but one patient (547) two mutated *CRB1* alleles were found.

The three novel missense mutations are in the conserved domains of the CRB1 protein. The p.Ser740Phe exchange replaces a highly conserved serine in the second laminin AG-like domain, the p.Tyr1198Cys mutation replaces a conserved tyrosine with a cysteine in the 16th calcium binding EGF-like domain and the p.Cys1223Ser is a replacement of a conserved cysteine with a serine in the 17th calcium binding EGF-

like domain (Figure 1). The in-frame deletion p.Asn789del is also located in the second laminin AG-like domain. Other novel mutations (p.Cys423*, p.Leu655Trpfs*10, p.Ser1220Asnfs*62) result in premature stop codons, which most likely lead to nonsense mediated decay (Chang, et al., 2007) and therefore these alleles are considered as null alleles. Five novel mutations are within exons 7 and 9, which are the most frequently mutated (Figure 1).

Clinical Characterisation of Patients with *CRB1* Mutations

Clinical findings of French patients with *CRB1* mutations are summarized in Tables 2 and 3. The average age at time of diagnosis was 17. Visual acuity was decreased in all patients ranging from 20/50 to light perception with no clear correlation with age or duration of the disease. Hyperopia was noted for 6/11 patients including three for whom spherical equivalent was equal or above +5 diopters. Night blindness was present in all patients but three, for whom a decrease of central vision and photophobia dominated. None of the patients had nystagmus. Most patients (9/11) had a clear lens; in the remaining two, one had undergone cataract surgery and one had significant lens opacities. These two patients were over 40 years of age. Two patterns of fundus pigmentary changes were present in this cohort: 7/11 had typical bone spicule-shaped pigment migration within the peripheral retina whereas 4/11 had widespread clumped pigmentary changes of nummular appearance at the level of the retinal pigment epithelium (Figure 2). Clumped pigmentation is therefore highly suggestive of *CRB1* mutations but it is not specific since it has also been associated with mutations in *NR2E3* (Schorderet and Escher, 2009; Sharon, et al., 2003), *NRL* (Nishiguchi, et al., 2004) or *TULP1* (Mataftsi, et al., 2007). None of the patients displayed preservation of the para-arteriolar retinal pigment epithelium as previously

described in association with *CRB1* mutations (Bernal, et al., 2003; den Hollander, et al., 2004; den Hollander, et al., 1999; Heckenlively, 1982; Henderson, et al., 2010; Khaliq, et al., 2003; Simonelli, et al., 2007; Yzer, et al., 2006b). In addition, none of the patients displayed Coats-like changes in the periphery. All patients had macular involvement. Six of the patients displayed cystoid macular edema whereas the other five had macular thinning with loss of the outer retinal layers and corresponding loss of autofluorescence (Figure 2). Color vision was normal in four patients or showed either tritan deficit or a dyschromatopsia with no clear axis when visual acuity allowed color vision testing. Full field electroretinogram showed severe generalized retinal dysfunction with no detectable responses in all patients except three for whom some residual rod and cone function was detectable. Among those three, the best responses on ERG were obtained in the youngest patients. Residual responses on ERG were correlated with better preservation of the visual field.

All patients displayed severe retinal involvement with early macular changes, half of them had cystoid macular edema, a higher percentage than the usually reported prevalence of about 30% in overall RP (Hajali, et al., 2008). This higher prevalence could at least be in part related to vascular abnormalities with Coats-like changes encountered in patients with *CRB1* mutations (Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001a; Henderson, et al., 2010; Yzer, et al., 2006b). Alternatively, these changes could be related to abnormal laminar structure associated with *CRB1*-mutations (Jacobson, et al., 2003). None of our patients developed Coats-like changes or para-arteriolar retinal pigment epithelium suggesting that these changes are not consistant in *CRB1*-related RP (Lotery, et al., 2001b). Four subjects displayed clumped retinopathies reinforcing that *CRB1* should be considered as a

potential causal gene for this specific phenotype along with *NR2E3* (Sharon, et al., 2003) or *NRL* (Nishiguchi, et al., 2004).

CRB1 Variants and Their Classification

Over 240 patients with *CRB1* mutations and more than 150 gene variants have been described in the literature (Benayoun, et al., 2009; 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; den Hollander, et al., 2007; den Hollander, et al., 1999; Galvin, et al., 2005; Gerber, et al., 2002; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Khaliq, et al., 2003; Li, et al., 2011; Lotery, et al., 2001a; Lotery, et al., 2001b; Riveiro-Alvarez, et al., 2008; Seong, et al., 2008; Simonelli, et al., 2007; Tosi, et al., 2009; Vallespin, et al., 2007; Yzer, et al., 2006a; Yzer, et al., 2006b; Zenteno, et al., 2011; Zernant, et al., 2005). The most frequently occurring of the known mutations is the p.Cys948Tyr in exon 9 (96 alleles reported, 24% of known *CRB1* mutations) (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; 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., 2007; Yzer, et al., 2006a; Zernant, et al., 2005). In general most of the mutations are in exons 9 (41%) and 7 (27%), therefore as a screening strategy these exons can be tested in the first instance (Figure 1, Supp. Table S1). Exons 7 and 9 encode second and third laminin AG-like domains respectively, implying that these domains are particularly important for CRB1 function. Missense mutations constitute

66% of all known mutations, the remaining being frameshift, truncation and splice site mutations.

We have attempted to classify all the reported mutations in three groups: 1) likely pathogenic, 2) unclassified variants, 3) unlikely pathogenic. This classification was based on the genetic data available from the literature, amino acid conservation and bioinformatic pathogenicity prediction tools (Supp. Tables S1-S3). An important criterion was the presence of two mutant alleles and co-segregation in the family. Approximately 30% of cases were reported with only one mutant allele, assuming that the second mutation is within the intronic region. For these patients however, one cannot exclude the possibility that there is another molecular cause of the pathology. The lack of the second mutant *CRB1* allele is sometimes explained by a digenic inheritance, however so far it has not been proven by co-segregation analysis (Li, et al., 2011; Vallespin, et al., 2007).

Pathogenicity is easier to asses in deletions and frameshift variants than in the case of missense changes, hence the importance of the bioinformatic analysis of the pathogenicity, amino acid conservation and functional analysis of the variants. On this basis we have not considered two changes identified in our cohort as pathogenic (p.Gly959Ser and p.Ala1354Thr) (den Hollander, et al., 2004; den Hollander, et al., 2001a)). The respective patients did not carry a second *CRB1* mutation and we did not consider the p.Gly959Ser and p.Ala1354Thr substitutions as likely pathogenic, based on poor conservation of the residues and low pathogenicity predictions using online bioinformatic tools: PolyPhen-2 and SIFT (Supp. Tables S2 and S3). One report suggests involvement of *CRB1* in autosomal dominant pigmented paravenous chorioretinal atrophy (McKay, et al., 2005), though the reported mutation

p.Val162Met has a questionable pathogenicity, since valine is not conserved and methionine is present in this position in other mammals (Supp. Table S2).

Prevalence

In the investigated cohort, at least 2.5% of arRP patients carry *CRB1* gene defects, which lies within the previously published range of 0-6.5% (Bernal, et al., 2003; den Hollander, et al., 2004; Vallespin, et al., 2007), or 2.7% after cohort averaging (Table 4). The high preponderance of novel *CRB1* mutations in our cohort suggests, however, that probably more arRP patients carry *CRB1* pathogenic defects, which are novel and therefore undetectable by arRP microarray. Much higher prevalence is observed in LCA/EORD cohorts and RP with additional features like PPRPE and retinal telangiectasia, representing 10.1%, 74.1%, 53.3% respectively in averaged cohorts (Table 4) (Bernal, et al., 2003; Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001a; den Hollander, et al., 2007; den Hollander, et al., 1999; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Lotery, et al., 2001a; Seong, et al., 2008; Simonelli, et al., 2007; Vallespin, et al., 2007; Walia, et al.).

Genotype-Phenotype Correlation

We were not able to establish a clear genotype/phenotype correlation for our cohort, which might be due to the small number of patients with *CRB1* mutations and their variable phenotype. In addition, the nature of existing published data makes it difficult to correlate the recurring *CRB1* mutations with different phenotypes for a number of reasons. First, the phenotyping of patients is complex and distinguishing between early-onset RP and LCA is often arbitrary and depends on the guidelines of a

particular clinical center. Second, precise clinical data is often omitted in the publications and therefore it is difficult to adjust for these diagnostic differences in a cross-paper analysis. Despite these inconsistencies, we attempted to analyse data from previous reports in order to find the relationship between the *CRB1* variants and the clinical features of respective patients. In this meta-analysis we used 171 patients, who carried two likely pathogenic mutations in trans (Benayoun, et al., 2009; 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; 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; Khaliq, et al., 2003; Li, et al., 2011; Lotery, et al., 2001a; Lotery, et al., 2001b; McKibbin, et al., 2010; Riveiro-Alvarez, et al., 2008; Seong, et al., 2008; Simonelli, et al., 2007; Tosi, et al., 2009; Vallespin, et al., 2007; Yzer, et al., 2006a). Combination of two mutant alleles was analysed in relation to clinical characteristics of the published cases. Based on the reports we distinguished the following phenotypes: LCA, early onset retinal degeneration (EORD), RP, presence of PPRPE and Coats-like vasculopathy. The mutations were classed as null mutations (all mutations leading to a premature stop codon) or as variants leading to an altered protein (missense and in frame deletions). The likely pathogenic mutations were plotted on a graph, where affected codons on allele 1 and allele 2 served as coordinates (codon 0 was assigned to null mutations). The results show that we cannot assign a specific allele combination to a particular phenotype, e.g. homozygous null alleles or homozygous p.Cys948Tyr alleles are found in LCA, EORD and RP patients (Figure 3 A). Null alleles are however more frequent in LCA cohorts (Figure 3 B) as previously suggested (den Hollander, et al., 2004). The presence/absence of PPRPE or Coats-like vasculopathy did not reveal a

particular mutation pattern (Figure 3 C). These findings suggest the involvement of additional modifying factors (genetic and/or environmental), which are responsible for the modulation of the phenotype in patients harboring *CRB1* mutations.

Future Directions

The above analysis of the phenotype-genotype correlation suggests that the disease severities associated with *CRB1* mutations are in fact a continuum of the same clinical entity with possible additional modifying factors influencing disease onset and progression. There is increasing evidence of the involvement of multiple alleles in the patient's phenotype, as has been shown for the Bardet-Biedl patients (Katsanis, et al., 2001) and more recently for a *PRPH2*-associated macular dystrophy family, where the phenotype has been modulated by additional heterozygous mutations in *ABCA4* (MIM#: 601691) and *ROM1* (MIM#: 180721) (Poloschek, et al., 2010). It is likely that the new next generation sequencing (NGS) technology will help to shed light on the potential genetic modifiers that influence disease phenotype. One has, however, to analyse the data with caution since NGS will reveal large numbers of polymorphic changes, which do not modulate the disease. The potential new modifying changes will have to be confirmed by appropriate genetic and functional analysis. The certainty of the molecular cause of a disease is particularly important in the era of gene therapy trials. Genetic treatment of recessive disorders should not be undertaken before obtaining proof that both alleles of a given gene are dysfunctional. In-depth genetic analysis, as presented here, is necessary to provide a basis for conducting such therapies.

Acknowledgments

The authors would like to thank patients and families for participation in this study, Dominique Santiard-Baron, Christine Chaumeil and clinical staff for their help in clinical data and DNA collection, Sandro Banfi, Robert Henderson and Qingjiong Zhang for additional information on genotype-phenotype correlations of previously published mutations and Robert Gillan for help with the manuscript. The project was financially supported by the Foundation Fighting Blindness (I.A. FFB Grant No: CD-CL-0808-0466-CHNO and the CIC503 recognized as an FFB center, FFB Grant No: C-CMM-0907-0428-INSERM04), Agence Nationale de la Recherche (SSB), Fondation Voir et Entendre (CZ), GIS-maladies rares (CZ), Ville de Paris and Région Ile de France, National Institutes of Health (USA) (KB NIH, Grant No: 1R01EY020902 - 01A1).

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 and others. 2010. EYS is a major gene for rod-cone dystrophies in France. Hum Mutat 31:E1406-35.

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.

Cahill M, O'Keefe M, Acheson R, Mulvihill A, Wallace D, Mooney D. 2001. Classification of the spectrum of Coats' disease as subtypes of idiopathic retinal telangiectasis with exudation. Acta Ophthalmol Scand 79:596-602.

Chang YF, Imam JS, Wilkinson MF. 2007. The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem 76:51-74.

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 and others. 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 and others. 2004. CRB1 mutation spectrum in inherited retinal dystrophies. Hum Mutat 24:355-69.

den Hollander AI, Ghiani M, de Kok YJ, Wijnholds J, Ballabio A, Cremers FP, Broccoli V. 2002. Isolation of Crb1, a mouse homologue of Drosophila crumbs, and analysis of its expression pattern in eye and brain. Mech Dev 110:203-7.

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 and others. 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 and others. 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, Roepman R, Koenekoop RK, Cremers FP. 2008. Leber congenital amaurosis: genes, proteins and disease mechanisms. Prog Retin Eye Res 27:391-419.

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 and others. 1999. Mutations in a human homologue of Drosophila crumbs cause retinitis pigmentosa (RP12). Nat Genet 23:217-21.

Franceschetti A, Dieterle P. 1954. [Diagnostic and prognostic importance of the electroretinogram in tapetoretinal degeneration with reduction of the visual field and hemeralopia]. Confin Neurol 14:184-6.

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 and others. 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.

Gosens I, den Hollander AI, Cremers FP, Roepman R. 2008. Composition and function of the Crumbs protein complex in the mammalian retina. Exp Eye Res 86:713-26.

Hajali M, Fishman GA, Anderson RJ. 2008. The prevalence of cystoid macular oedema in retinitis pigmentosa patients determined by optical coherence tomography. Br J Ophthalmol 92:1065-8.

Hanein S, Perrault I, Gerber S, Tanguy G, Barbet F, Ducroq D, Calvas P, Dollfus H, Hamel C, Lopponen T and others. 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.

Heckenlively JR. 1982. Preserved para-arteriole retinal pigment epithelium (PPRPE) in retinitis pigmentosa. Br J Ophthalmol 66:26-30.

Henderson RH, Mackay DS, Li Z, Moradi P, Sergouniotis P, Russell-Eggitt I, Thompson DA, Robson AG, Holder GE, Webster AR and others. 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 and others. 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.

Katsanis N, Ansley SJ, Badano JL, Eichers ER, Lewis RA, Hoskins BE, Scambler PJ, Davidson WS, Beales PL, Lupski JR. 2001. Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. Science 293:2256-9.

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.

Leber T. 1869. Ueber Retinitis pigmentosa und angeborene Amaurose. Graefe's Archive For Clinical And Experimental Ophthalmology 15:1-25.

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 and others. 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.

Mataftsi A, Schorderet DF, Chachoua L, Boussalah M, Nouri MT, Barthelmes D, Borruat FX, Munier FL. 2007. Novel TULP1 mutation causing leber congenital amaurosis or early onset retinal degeneration. Invest Ophthalmol Vis Sci 48:5160-7.

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.

McMahon TT, Kim LS, Fishman GA, Stone EM, Zhao XC, Yee RW, Malicki J. 2009. CRB1 gene mutations are associated with keratoconus in patients with leber congenital amaurosis. Invest Ophthalmol Vis Sci 50:3185-7.

Mehalow AK, Kameya S, Smith RS, Hawes NL, Denegre JM, Young JA, Bechtold L, Haider NB, Tepass U, Heckenlively JR and others. 2003. CRB1 is essential for external limiting membrane integrity and photoreceptor morphogenesis in the mammalian retina. Hum Mol Genet 12:2179-89.

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

Nishiguchi KM, Friedman JS, Sandberg MA, Swaroop A, Berson EL, Dryja TP. 2004. Recessive NRL mutations in patients with clumped pigmentary retinal degeneration and relative preservation of blue cone function. Proc Natl Acad Sci U S A 101:17819-24.

Pellikka M, Tanentzapf G, Pinto M, Smith C, McGlade CJ, Ready DF, Tepass U. 2002. Crumbs, the Drosophila homologue of human CRB1/RP12, is essential for photoreceptor morphogenesis. Nature 416:143-9.

Poloschek CM, Bach M, Lagreze WA, Glaus E, Lemke JR, Berger W, Neidhardt J. 2010. ABCA4 and ROM1: implications for modification of the PRPH2-associated macular dystrophy phenotype. Invest Ophthalmol Vis Sci 51:4253-65.

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.

Schorderet DF, Escher P. 2009. NR2E3 mutations in enhanced S-cone sensitivity syndrome (ESCS), Goldmann-Favre syndrome (GFS), clumped pigmentary retinal degeneration (CPRD), and retinitis pigmentosa (RP). Hum Mutat 30:1475-85.

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.

Sharon D, Sandberg MA, Caruso RC, Berson EL, Dryja TP. 2003. Shared mutations in NR2E3 in enhanced S-cone syndrome, Goldmann-Favre syndrome, and many cases of clumped pigmentary retinal degeneration. Arch Ophthalmol 121:1316-23.

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

Tepass U, Theres C, Knust E. 1990. crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila epithelial cells and required for organization of epithelia. Cell 61:787-99.

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

van de Pavert SA, Kantardzhieva A, Malysheva A, Meuleman J, Versteeg I, Levelt C, Klooster J, Geiger S, Seeliger MW, Rashbass P and others. 2004. Crumbs homologue 1 is required for maintenance of photoreceptor cell polarization and adhesion during light exposure. J Cell Sci 117:4169-77.

van de Pavert SA, Meuleman J, Malysheva A, Aartsen WM, Versteeg I, Tonagel F, Kamphuis W, McCabe CJ, Seeliger MW, Wijnholds J. 2007. A single amino acid substitution (Cys249Trp) in Crb1 causes retinal degeneration and deregulates expression of pituitary tumor transforming gene Pttg1. J Neurosci 27:564-73.

Walia S, Fishman GA, Jacobson SG, Aleman TS, Koenekoop RK, Traboulsi EI, Weleber RG, Pennesi ME, Heon E, Drack A and others. 2010. Visual acuity in patients with Leber's congenital amaurosis and early childhood-onset retinitis pigmentosa. Ophthalmology 117:1190-8.

Yzer S, Fishman GA, Racine J, Al-Zuhaibi S, Chakor H, Dorfman A, Szlyk J, Lachapelle P, van den Born LI, Allikmets R and others. 2006a. CRB1 heterozygotes with regional retinal dysfunction: implications for genetic testing of leber congenital amaurosis. Invest Ophthalmol Vis Sci 47:3736-44.

Yzer S, Leroy BP, De Baere E, de Ravel TJ, Zonneveld MN, Voesenek K, Kellner U, Ciriano JP, de Faber JT, Rohrschneider K and others. 2006b. 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 and others. 2005. Genotyping microarray (disease chip) for Leber congenital amaurosis: detection of modifier alleles. Invest Ophthalmol Vis Sci 46:3052-9.

Figure Legends

Figure 1. Distribution of *CRB1* mutations in the gene and protein. A) Nucleotide numbering is based on cDNA sequence of *CRB1* (Ref. NM_201253.2) where A of the ATG initiation codon is 1. The stop and frameshift mutations are indicated above the structure of the gene and the position of the missense mutations are drawn in relation to protein domains. The novel mutations are indicated in red. B) The structures of EGF-like and Ca^{++} binding EGF-like domains with indications of conserved residues and recurrent mutations. The highly conserved cysteine residues are in black, the conserved residues between both domains are in grey and the conserved amino acids specific to the Ca^{2+} binding domain are in blue. C) Evolutionary conservation of the likely pathogenic CRB1 residue changes identified in this work.

Figure 2. Fundus color photographs and Optical Coherence Tomography (OCT). A) Color fundus photograph of the left eye of 3969 showing nummular pigmentary migration in the mid periphery in addition to pigmentary changes within the macula. B) Vertical scan OCT of the left eye of 3969 showing cystic changes in the macular region. C) Color fundus photograph of the right eye of 547 showing bone spicules pigmentary migration in the periphery in addition to atrophic changes within the macula. D) Vertical scan OCT of the right eye of 547 showing atrophic changes in the macular region after resolution of episodes of cystoid changes.

Figure 3. Genotype-phenotype correlation of patients with *CRB1* mutations. A) Distribution of *CRB1* mutations in LCA, EORD and RP. XY axes represent allele 1 and 2 of the patients, the affected codons serve as xy coordinates, null allele coordinate is designated as 0. The size of the circles is proportional to the number of the *CRB1* patients with a given genotype. B) Frequency of null and missense allele combinations in LCA, EORD and RP patients. C) Distribution of *CRB1* mutations in patients with/without additional features: PPRPE and Coats-like vasculopathy.

Table 1. Patients with *CRB1* mutations identified in this study

Patient number	Family	Allele 1			Allele 2		
		Exon	Nucleotide change	Protein change	Exon	Nucleotide change	Protein change
229	159	2	c.613_619del	p.Ile205Aspfs*13	**7**	**c.2365_2367del AAT**	**p.Asn789del**
53	No family members	**6**	**c.1269C>A**	**p.Cys423***	7	c.2506C>A	p.Pro836Thr
368	249	6	c.1750G>T	p.Asp584Tyr	7	c.2506C>A	p.Pro836Thr
547	372	**6**	**c.1963delC**	**p.Leu655Trpfs*10**		?	
4240[a]	2025	**7**	**c.2219C>T**	**p.Ser740Phe**	**7**	**c.2219C>T**	**p.Ser740Phe**
54	39	7	c.2222T>C	p.M741T	**9**	**c.3593A>G**	**p.Tyr1198Cys**
3969	No family members	7	c.2506C>A	p.Pro836Thr	7	c.2506C>A	p.Pro836Thr
409	281	9	c.2843G>A	p.Cys948Tyr	**9**	**c.3668G>C**	**p.Cys1223Ser**
1183[b]	709	**9**	**c.3659_3660delinsA**	**p.Ser1220Asnfs*62**	**9**	**c.3659_3660del insA**	**p.Ser1220Asnfs*62**
1731	1008	9	c.2843G>A	p.Cys948Tyr	9	c.2843G>A	p.Cys948Tyr
3144	1302	9	c.2843G>A	p.Cys948Tyr	7	c.3307G>A	p.Gly1103Arg

[a] mutation in this patient was identified by NGS

[b] mutation in this patient was found through homozygosity mapping

novel mutations are in bold

Table 2: Clinical data

Patient	Age at time of testing	Age at time of diagnosis	Sex	Relevant medical and ophthalmology history	Family history	Symptoms	BCVA OD/OS Refraction	Lens	Fundus examination	OCT	FAF
53	27	20	M	none	From Ivory Coast, 10 brothers and sisters, 1 sister affected	Night blindness at 6 then photophobia then decreased vision	LP 20/500 +2(-1.50)60° +1.75(-1.5)125°	Clear	Widespread clumped pigment migration with no pale optic disc or narrowed retinal vessels	Macular thinning with loss of ONL	Loss of AF at the posterior pole and periphery
54	41	25	F	none	From French descent One affected brother	Night blindness	20/640 20/100 Prior to lens surgery: +5.50(-1)5° +5.50(-1)165°	IOL	Peripheral RPE changes with bone spicules, perifoveal atrophy, pale optic disc, narrowing of retinal vessels	Thinning of the ONL within the macular region	Loss of AF in the perifoveal region and outside the vascular arcades
229	29	20	F	none	From French descent	Night blindness	20/80 20/50 +2(-0.75)5° +2.50(-1.50)5°	Clear	Peripheral RPE changes, little bone spicules, no pale optic disc or narrowed retinal vessels, CME	CME, thinning of ONL	Patchy loss of AF in the periphery; foveal modification of AF due to the CME
368	13	12	F	Seizure in infancy	From Turkish descent maternal grand-mother said to be blind	photophobia	20/80 20/63 +6.50(-1.25)160° +6.50(-1)7°	Clear	Peripheral RPE changes with bone spicules, perifoveal atrophy, pale optic disc, narrowing of retinal vessels, CME	CME with relative preservation of foveal architecture	Patchy loss of AF outside the vascular arcades, foveal AF changes due to CME
409	43	Teenage years	F	none	From Italian descent	Night blindness then photophobia	20/160 20/100 Plano Plano	Clear	Peripheral bone spicules with perifoveal atrophy	Thinning of the ONL	Loss of AF outside the vascular arcades and in the perifoveal area

547	57	39	M	Recurrent anterior uveitis, which delayed the diagnosis of RP	From French descent, no family history of RP	Night blindness then photophobia and decreased vision	20/80 20/63 +0.25(-0.50)110° -2(-1.25)65°	Bilateral nuclear cataract	Peripheral bone spicules with CME	Bilateral CME, perifoveal thinning	Loss of AF in the perifoveal region and outside the vascular arcades
1183	38	15	F	none	From Tunisian descent; consanguinity among parents	Night blindness and photophobia	20/640 20/640 Emetropia	Clear	Widespread clumped pigment migration with no pale optic disc or narrowed blood vessels; OD asteroides hyaloids	Macular thinning with loss of ONL	Loss of AF at the posterior pole and periphery
1731	23	17	M	Deafness since age 9	From Spanish descent; parents first cousins; one brother affected	Low vision since early childhood	HM 20/80 Emetropia	Clear	Widespread clumped pigment migration with relative sparing of the macula, with no pale optic disc or narrowed blood vessels	Macular thinning with loss of ONL	Loss of AF at the posterior pole and periphery
3144	20	9	F	none	From French descent	Night blindness since early childhood	20/80 20/80 +9(-1.50)170° +7.50	Clear	Some RPE changes in the periphery, normal disc color and no narrowing of blood vessels; CME	CME with relatively spared foveal structure	Patchy loss of AF outside the vascular arcades, foveal AF changes due to CME
3969	28	12	F	none	From Mali	Night blindness then photophobia	20/125 20/320 +0.50(-1.50)90° +1.75(-1.25)95°	Clear	Widespread clumped pigment migration in the posterior pole and periphery CME	CME Thinning of ONL	Diffuse patchy loss of AF within the posterior pole and periphery
4240	7	6	M	none	One sister affected, from Turkish descent	Decreased vision	20/63 20/80 -1.50(-1.50)10° -2(-0.75)180°	Clear	Moderate RPE changes in the periphery CME	CME with relatively spared parafoveal	Patchy loss of AF outside the vascular arcade, normal AF

										structure	within posterior pole except AF modification due to CME in the fovea

BCVA: best corrected visual acuity; CME: cystoid macular edema; ND: not detectable; FAF: Fundus Autofluorescence; OD: Oculis dextra (right eye); OS: Oculis Sinistra (left eye); IOL: intra ocular lens; CF: counting fingers; HM: hand motion; LP: light perception; RPE: retinal pigment epithelium; RP: retinitis pigmentosa; OHT: ocular hypertension; ONL: Outer Nuclear Layer

Table 3: Function data

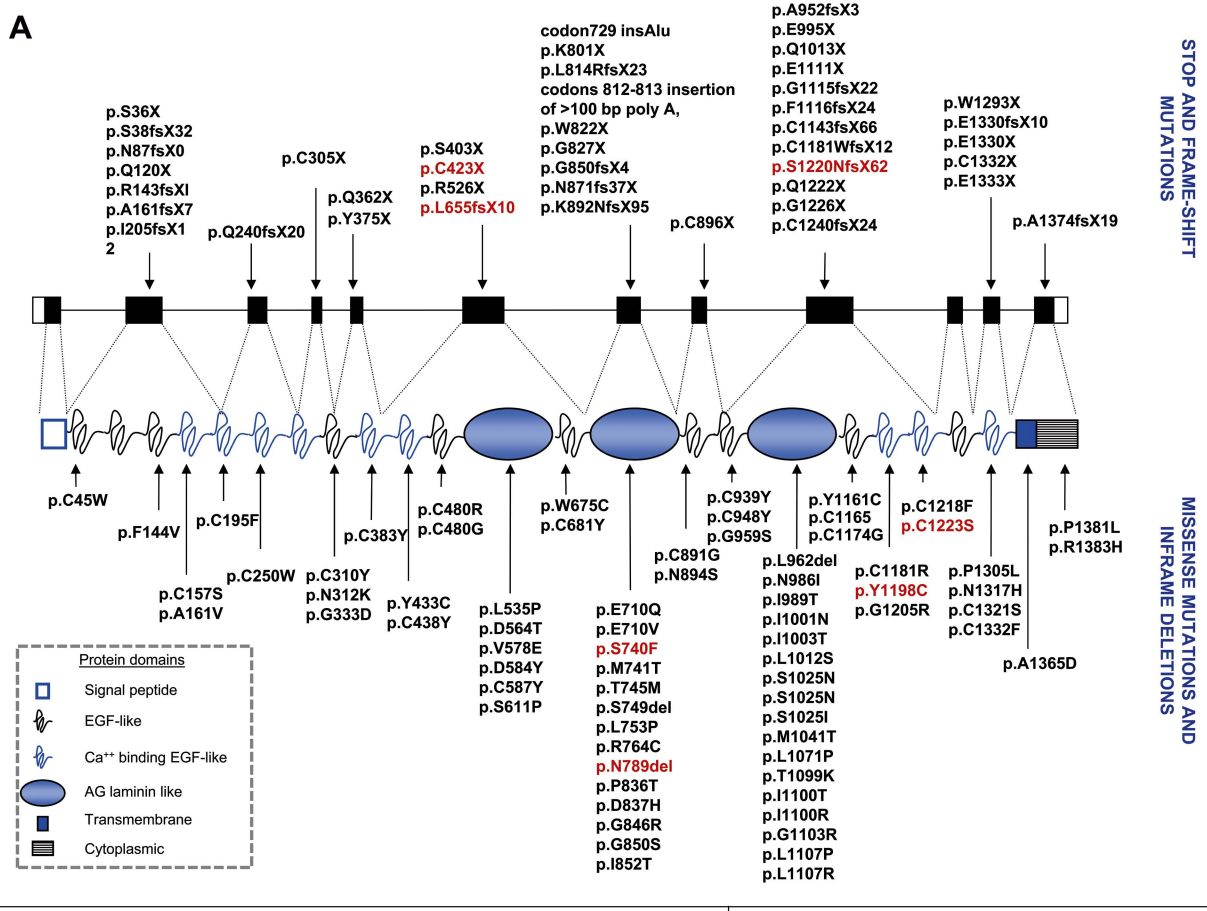
Patient	Colour vision (15 saturated Hue)	Binocular Goldman visual field, III4 isopter	Full field ERG	Multifocal ERG
53	NP	Inf to 5°	ND	ND
54	Dyschromatopsia without axis	Inf to 5°	ND	ND
229	Normal	40 central degree with 2 peripheral island of perception	ND	ND
368	Normal	120° horizontally, 60° vertically with relative central annular scotoma	Residual responses consistent with severe rod-cone dysfunction	Residual responses to central hexagones
409	Dyschromatopsia without axis	100° horizontally, 60° vertically with annular scotoma	Residual cone responses	ND
547	Bilateral tritaonopia	20 central degrees both horizontally and vertically	ND	ND
1183	NP	Inf to 5°	ND	ND
1731	OD NP, OS tritaonopia	5 central degrees	ND	ND
03144	Normal	20 central degrees both horizontally and vertically	ND	ND
3969	Dyschromatopsia without axis	20 central degree with 2 peripheral island of perception	ND	ND
4240	Normal	130° vertically and 110° horizontally	30% decreased scotopic responses with photopic responses at the lower limit of normal	Decreased responses to central hexagones

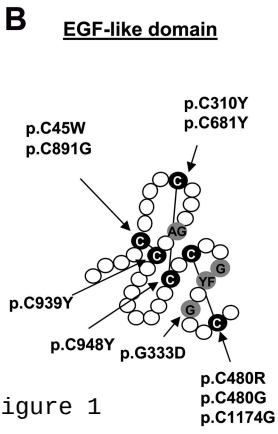
NP: not performed; ND: not detectable

Table 4. Average prevalence of *CRB1* mutations in retinal dystrophy patients in published reports

Dystrophy	Prevalence*	Patients with two CRB1 alleles	Patients with one CRB1 allele	Added cohort size	References
LCA/EORD	10.1%	109	57	1645	(Bernal, et al., 2003; Coppieters, et al., 2010; den Hollander, et al., 2004; den Hollander, et al., 2001; den Hollander, et al., 2007; den Hollander, et al., 1999; Hanein, et al., 2004; Henderson, et al., 2010; Henderson, et al., 2007; Li, et al., 2011; Lotery, et al., 2001; Seong, et al., 2008; Simonelli, et al., 2007; Vallespin, et al., 2007; Walia, et al.)
RP	2.7%	4	5	335	(Bernal, et al., 2003; den Hollander, et al., 2004; Vallespin, et al., 2007)
RP+PPRPE	74.1%	18	2	27	(den Hollander, et al., 2004; den Hollander, et al., 1999)
RP+ret telangiectasia	53.3%	8	8	30	(den Hollander, et al., 2004; den Hollander, et al., 2001; Henderson, et al., 2010)
Classic Coats disease	0.0%	0	0	18	(den Hollander, et al., 2004)

* The average prevalence was calculated on the basis of all the published reports indicating phenotypes of patients with *CRB1* mutations and the size of screened cohorts.





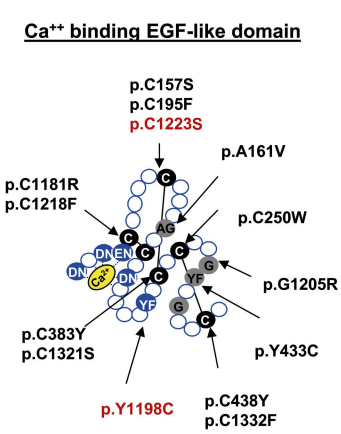


Figure 1

C

	p.S740F	p.N789del	p.Y1198C	p.C1223S
Human	DTISLSMFVRT	VKFVLNDGNVH	RVAAYHCTCE	QSHQCANGAT
Chimp	DTISLSMFVRT	VKFVLNDGNVH	RVAAYHCTCE	QSHQCANGAT
Rhesus	DTVSLSMFVQT	VKFVLNDGNVH	RVAAYHCTCE	QSHQCANGAT
Tarsier	EKFSLSMFVRT	VKFILSDGNVH	RVAAYYCRCE	QNHHCANGAT
Mouse lemur	ENISFSMLVRT	G=FAISDGH=H	AAAAYLCRCE	QRHQCANGAT
Bushbaby	ENFSLSMFVRT	VEFVISDGNFH	RVAAYHCGCE	RSHQCANGAT
Tree shrew	ENVSLSMFVRT	TEFLLSDGNAH	GVAAYHCRCE	RSHRCANGAT
Mouse	QNFSLSMFVRT	VNFFLSDGNVH	GVAAYHCRCE	KSHQCANGAT
Squirrel	QNFSLSMFVRT	VEFVLSDGNVH	RVAAYHCRCE	QNHQCANGAT
Rabbit	ENFSISMFIRT	VKFVLSDGNVH	QLAAYHCRCE	QGHQCANGAT
Cow	EDLTLSMFVRT	VTFVLSDGNVR	GGTAYHCRCE	QHHQCANGAT
Horse	ENFTLSMFVRT	VKFVLSDGNIH	RLAAYHCKCE	RSHQCANGAT
Cat	DNFTISMFVRT	VK==========	RVAAYHCRCE	QSHQCANGAT
Dog	ENVTISMFVRT	AKFVLSDGNVH	RVAAYRCRCE	QNHQCANGAT
Hedgehog	ENFTLSMFVRT	AEFVLSDGNVH	GVAAYHCRCE	QSHFCANGAA
Elephant	ENFTLSMFVRT	VRFVLNDGNAH	RVAAYHCTCE	QSHQCANGAT
Sloth	ENFTLSMFVRT	VKFFLSDGNSH	RVTAFHCRCE	QSHQCANGAT
Wallaby	EFVAVSMFVRT	GRSLLNDGNSH	RIAGYICICN	ENHQCANGAT
Opossum	EIVTLSMFVRT	GRSLLNDGNIH	RIAGYICKCN	ENHRCANGAT
Platypus	QNITLSMFVRT	GKTMLSDGNMH	RVASYVCVCE	RGHQCANGAT
Chicken	ENLNISMFVRT	GKHIINDGNFY	GINSYECICE	QGHQCANGAT
Lizard	ETATISMFVRT	GTLPVNDGNFY	RTASYTCICD	HGHLCANGAT
X. tropicalis	EEITISAFFRT	ANYTINNGNFH	TITGYLCKCQ	QDHQCATGAT
Tetraodon	PDLSVSFFLRT	VQSPVDDGTVH	RGPTFECACD	QNHRCANGGT
Stickleback	EDFSISLFLRT	DRSVIDDGEVH	KGPTYECTCE	EKHLCAHGGT
Zebrafish	RKHFISMFLRT	W*RVLNDGERH	QDLTYNCTCE	AGHRCANGAT

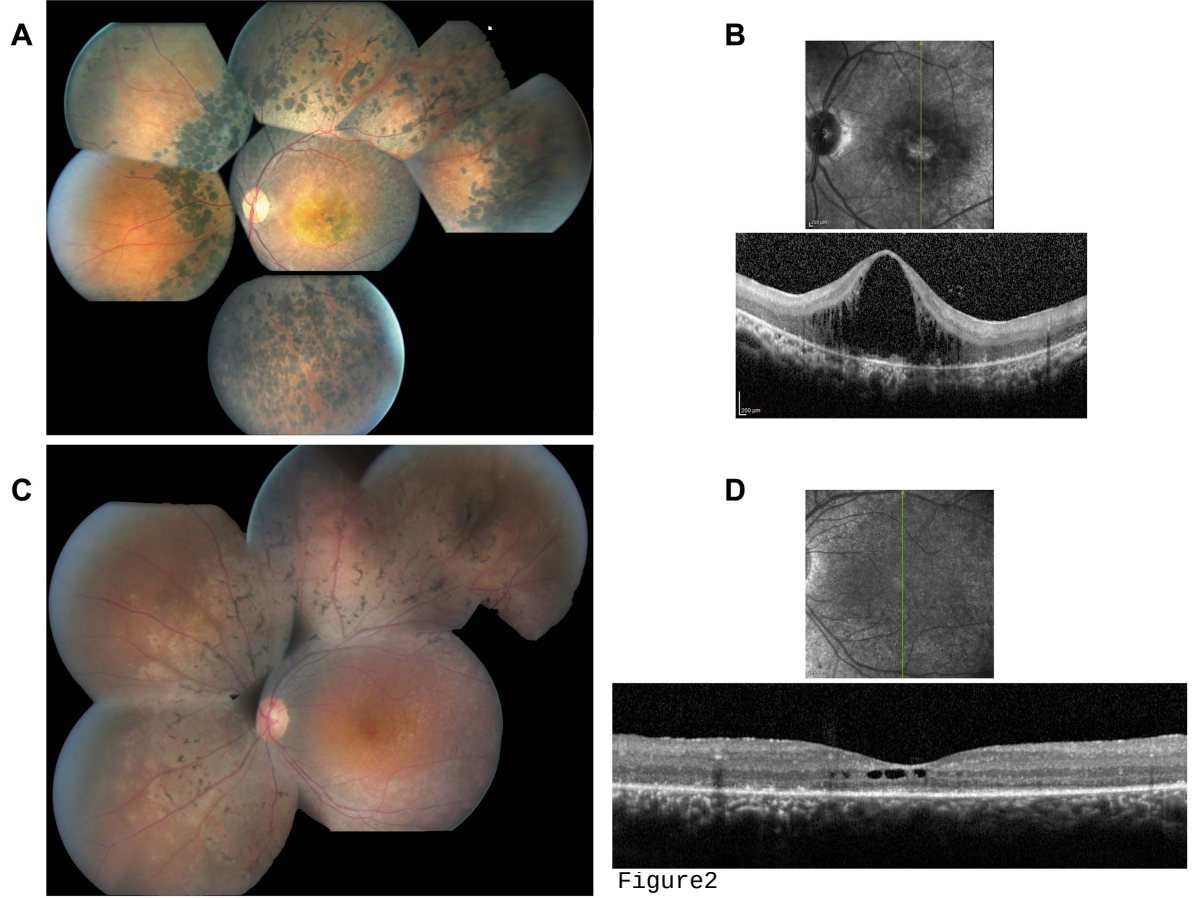
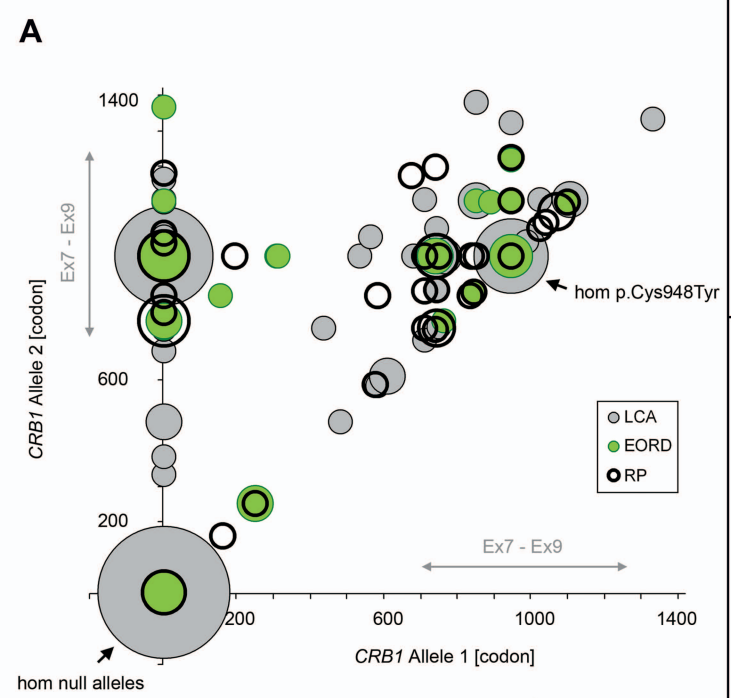
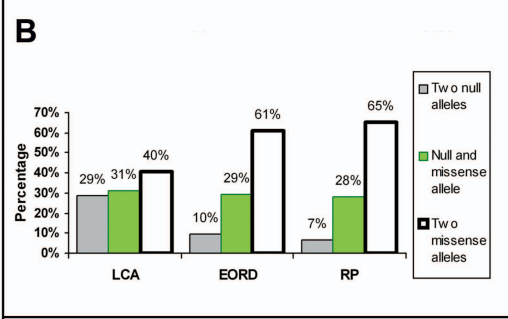


Figure2


A
1400
600
200
Ex7 - Ex9
CRB1 Allele 2 [codon]
hom p.Cys948Tyr
LCA
EORD
RP
Ex7 - Ex9
200
600
1000
1400
CRB1 Allele 1 [codon]
hom null alleles


B
70%
60%
50%
40%
30%
20%
10%
0%
Percentage
29%
31%
40%
10%
29%
61%
7%
28%
65%
LCA
EORD
RP
Two null alleles
Null and missense allele
Two missense alleles

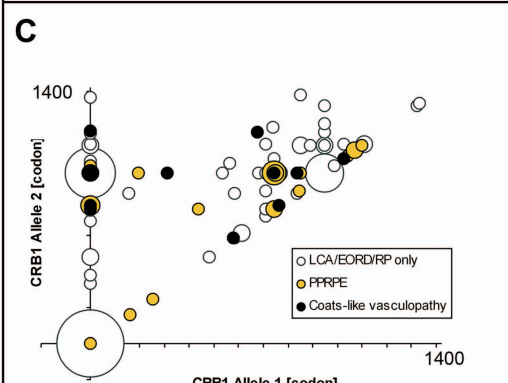

C
1400
CRB1 Allele 2 [codon]
LCA/EORD/RP only
PPRPE
Coats-like vasculopathy
1400
CRB1 Allele 1 [codon]

Figure 3

SUPPORTING MATERIAL

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

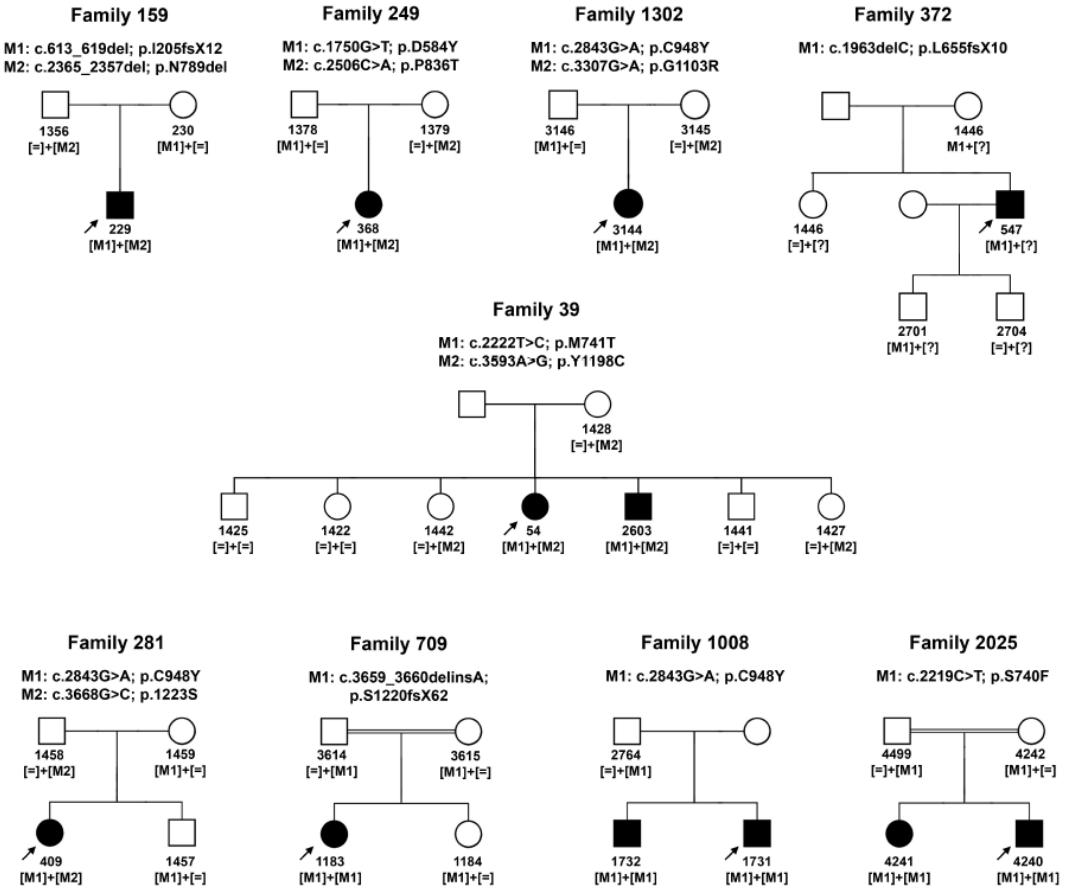
For Sanger sequencing, *CRB1* gene (*CRB1* RefSeq NM_201253) was PCR amplified in 15 fragments using oligonucl eotides flanking the exons and a polymerase (HotFire, Solis Biodyne, Estonia) in the presence of 1.5-2.0 mM $MgCl_2$

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.



Supplement 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.

Supplement 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	Probably Damaging (score	1	RP with PPRPE		(den Hollander, et al., 2004)

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 (considering 22 species, Val in Bushbaby, Mouse, Horse)	Tolerated (score 0.23)	Possibly Damaging (score 0.426)	2	LCA, RP		(Hanein, et al., 2004; Simonelli, et al., 2007)
7	c.2611_2613ins	p. Asn871Ilefs*38	LamAG 2	protein	-	-	1	LCA	originally it was	(Lotery, et al., 2001a)

					(score 0.00)	0.998)				
2	c.613_619del	p.Ile205Aspfs*13	EGF5	protein truncation, NMD	-	-	14	LCA EORD		(den Hollander, et al., 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)
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	Tolerated (score 0.08)	Probably Damaging	1	LCA		(Vallespin, et al., 2007b)

				(considering 26 species; Met in Squirrel)		(score 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)

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)
7	c.2258T>C	p.Leu753Pro	LamAG 2	Highly conserved up to Chicken	Tolerated (score 0.14)	Probably Damaging	1	LCA, ret talangiectas	p.Phe488Ser mutation on the	(Galvin, et al., 2005)

				(considering 20 species)		(score 0.994)		ia	second allele, which didn't co-segregate in the family	
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., 2005; Henderson, et al., 2010; Henderson, et al., 2007; Jacobson, et al., 2003; Simonelli, et al., 2007; Yzer, et al., 2006)

	T	or p.Ala872Cysfs*37		truncation, NMD					reported as insT in 871 codon	
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; 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 species)	Tolerated (score 0.37)	Probably Damaging (score 0.910)	2	LCA	considered as likely pathogenic due to cosegregation in the family	(Vallespin, et al., 2007b)
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	1	RP		(Henderson, et al., 2010)

						0.995)				
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)
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	Tolerated	Probably	2	LCA		(Hanein, et al., 2004; Henderson,

				(considering 25 species)	(score 0.24)	Damaging (score 0.997)				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 (score 0.00)	Probably Damaging (score 0.997)	2	LCA, RP ret telangiectasia		(Henderson, et al., 2010)
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**	**Affect**	**Probably**	**1**	**EORD**	**not found in**	**This study**

				conserved (considering 25 species, Phe in Sloth and Tetraodon)	**protein function (score 0.02)**	**Damaging (score 0.999)**			**378 control alleles**	
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)
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	Affect protein	Possibly Damaging	1	LCA	unknown second allele	(Lotery, et al., 2001a)

				(considering 24 species)	function (score 0.05)	(score 0.840)				
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 stretch to a polar Asp	(Henderson, et al., 2010)
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	Tolerated	Possibly	2	RP, RP	unknown	(Clark, et al., 2010; den

				conserved (considering 25 species, Gly in Mouse, Trp in Hedgehog)	(score 0.14)	Damaging (score 0.802)		with PPRPE	second allele	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; 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 Suppl. Methods.

Supplement 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	LamAG 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)
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)
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 Suppl. Methods.

Supplement 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)
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 Suppl. Methods.