A novel intronic mutation of *PDE6B* is a major cause of autosomal recessive retinitis pigmentosa among Caucasus Jews

Yasmin Tatour,[1] Jonathan Tamaiev,[1] Shamaly Shamaly,[2] Roberto Colombo,[3,4] Ephrat Bril,[5] Tom Rabinowitz,[6] Alona Yaakobi,[7] Eedy Mezer,[1,8] Rina Leibu,[8] Beatrice Tiosano,[1,7] Noam Shomron,[6] Itay Chowers,[5] Eyal Banin,[5] Dror Sharon,[5] Tamar Ben-Yosef[1]

[1]Ruth & Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; [2]Department of Ophthalmology, Bnai Zion Medical Center, Haifa, Israel; [3]Institute of Clinical Biochemistry, Faculty of Medicine, Catholic University of the Sacred Heart, Milan, Italy; [4]Center for the Study of Rare Hereditary Diseases, Niguarda Ca' Granda Metropolitan Hospital, Milan, Italy; [5]Department of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; [6]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [7]Department of Ophthalmology, Hillel Yaffe Medical Center, Hadera, Israel; [8]Alberto Moscona Department of Ophthalmology, Rambam Health Care Center, Haifa, Israel

Purpose: To identify the genetic basis for retinitis pigmentosa (RP) in a cohort of Jewish patients from Caucasia.
Methods: Patients underwent a detailed ophthalmic evaluation, including funduscopic examination, visual field testing, optical coherence tomography (OCT), and electrophysiological tests, electroretinography (ERG) and visual evoked potentials (VEP). Genetic analysis was performed with a combination of whole exome sequencing (WES) and Sanger sequencing. Bioinformatic analysis of the WES results was performed via a customized pipeline. Pathogenicity of the identified intronic variant was evaluated in silico using the web tool Human Splicing Finder, and in vitro, using a minigene-based splicing assay. Linkage disequilibrium (LD) analysis was used to demonstrate a founder effect, and the decay of LD over generations around the mutation in Caucasus Jewish chromosomes was modeled to estimate the age of the most recent common ancestor.
Results: In eight patients with RP from six unrelated families, all of Caucasus Jewish ancestry, we identified a novel homozygous intronic variant, located at position −9 of *PDE6B* intron 15. The c.1921–9C>G variant was predicted to generate a novel acceptor splice site, nine bases upstream of the original splice site of intron 15. In vitro splicing assay demonstrated that this novel acceptor splice site is used instead of the wild-type site, leading to an 8-bp insertion into exon 16, which is predicted to cause a frameshift. The presence of a common ancestral haplotype in mutation-bearing chromosomes was compatible with a founder effect.
Conclusions: The *PDE6B* c.1921–9C>G intronic mutation is a founder mutation that accounts for at least 40% (6/15 families) of autosomal recessive RP among Caucasus Jews. This result is highly important for molecular diagnosis, carrier screening, and genetic counseling in this population.

Retinitis pigmentosa (RP) is the most common form of inherited retinal dystrophy (IRD), with a worldwide prevalence of approximately 1 in 4,000. Classically, the first symptom of RP is night blindness, as rod photoreceptors are affected to a greater extent than cones. Disease progression is usually associated with a gradual restriction of the visual field. Ophthalmologic findings also include characteristic pigmentation of the midperipheral retina, attenuation of retinal arterioles, and optic disc pallor. In most patients, the disease is limited to the eye (nonsyndromic), with no extraocular manifestations [1]. RP is one of the most genetically heterogenous conditions in humans. It can be inherited in an autosomal recessive (AR), autosomal dominant (AD), or X-linked (XL) manner. To date, more than 85 genes have

been implicated in nonsyndromic RP, of which at least 62 are associated with an AR mode of inheritance (RetNet - Retinal Information Network).

One of the first genes identified to cause arRP was *PDE6B* (NM_000283; OMIM 180072) encoding the β subunit of rod photoreceptor cyclic GMP phosphodiesterase (PDE), a key enzyme of the visual phototransduction cascade [2,3]. Recessive biallelic mutations in *PDE6B* are a relatively common cause of arRP in various populations [4-6]. In addition, heterozygous dominant mutations in *PDE6B* are a rare cause of AD congenital stationary night blindness [7,8]. The *PDE6B* gene harbors 45 Mb on human chromosome 4p16.3, and is composed of 22 exons, which give rise to several splice isoforms, ranging in length from 2.7 to 3.4 Kb.

Caucasia is a region located at the border of Europe and Asia, situated between the Black Sea and the Caspian Sea. More than 50 different ethnic groups reside in this region, including Russians, Georgians, Azarians, and Armenians.

Correspondence to: Tamar Ben-Yosef, Rappaport Faculty of Medicine, Technion, P.O. Box 9649, Bat Galim, Haifa 31096, Israel; Phone: 972-4-829-5228; Fax: 972-4-829-5225; email: benyosef@technion.ac.il

Caucasus Jews (also known as "Mountain Jews") were defined as Jews who resided in North and East Caucasia, mainly in Azerbaijan, Chechnya, Dagestan, and the Kabardino Balkaria Republic and the Republic of North Ossetia-Alania, which are located in southern Russia. Between 1970 and 1990, most Caucasus Jews emigrated, mainly to Israel, but also to the United States and Russia [9]. Here, we report a novel intronic founder mutation of *PDE6B,* which affects splicing, and is a major cause of arRP among Caucasus Jews.

METHODS

Patients: A total of 17 Israeli RP patients were recruited for this study, including 8 females and 9 males, with an age range of 20-78 years old. Nine of them were recruited through Hadassah-Hebrew University Medical Center in Jerusalem, six through Rambam Health Care Center in Haifa, one through Hillel Yaffe Medical Center in Hadera, and one through Tel Aviv Sourasky Medical Center. The tenets of the Declaration of Helsinki were followed, the study adhered to the ARVO statement on human subjects, the study was approved by institutional review boards in participating Medical Centers, and written informed consent was obtained from all participants. All patients reported Caucasus Jewish ancestry on both parental sides.

Ophthalmic examination included measurement of best corrected visual acuity (BCVA), fundus examination after pupillary dilatation, fundus photography, spectral domain–optical coherence tomography (SD-OCT; Heidelberg Engineering, Heidelberg, Germany), widefield autofluorescence and pseudocolor fundus imaging (Optos, Dunfermline, UK), full-field electroretinography (ff-ERG), and measurement of flash visual evoked potentials (VEP). ffERG was conducted according to International Society for Clinical Electrophysiology of Vision (ISCEV) guidelines [10], using UTAS Bigshot electrophysiological system (LKC Technologies, Gaithersburg, MD), with bipolar Burian-Allen corneal electrodes (Hansen Ophthalmic Development, Coralville, IA).

DNA analyses: Venous blood samples were obtained using K3EDTA vacuette tubes (Greiner Bio-One, Kremsmunster, Austria) and kept at 4 ºC for up to 48 h. Genomic DNA was extracted from venous blood samples using a high-salt solution according to a standard protocol [11]. Whole exome sequencing (WES) of patient A-1 was performed at Theragen Etex Bio Institute (Suwon, Korea) using SureSelectXT Human All Exon V5 (Agilent Technologies, Santa Clara, CA) and the HiSeq 2500 Sequencing System (Illumina, San Diego, CA). Sequence reads were aligned to the reference human genome (GRCh37/hg19), and variants were called

via a customized pipeline as previously described [12]. For confirmation of the c.1921–9C>G mutation and screening of additional patients, specific primers were used to PCR-amplify *PDE6B* exon 16, including intron–exon boundaries (forward, 5′-GAG AGG CAC AGG CAG CCG AG-′3 and reverse, 5′-CCG TGG CGA TGA TGG CGA TG-′3). PCR was performed on 50 ng of genomic DNA in a 25 µl reaction volume in the presence of 5X Readymix (LAROVA GmbH, Teltow, Germany) and 10 pmol of each forward and reverse primers. Mutation screening was performed by direct sequencing with the Big Dye terminator cycle sequencing kit on an ABI 3130xl Genetic Analyzer (PE Applied Biosystems, Foster City, CA).

Splicing analysis: Evaluation of the identified intronic variant was performed using the web tool Human Splicing Finder [13]. In vitro evaluation was performed using a minigene-based assay. To create wild-type (wt) and mutant minigene constructs, a 750 bp DNA fragment harboring *PDE6B* exon 15, intron 15, and exon 16 was PCR-amplified from genomic DNA of patients and controls. The fragment was inserted into a mammalian expression vector harboring the Cytomegalovirus (CMV) promoter (pCMV-Script). Constructs were transfected into COS-7 cells, using the jetPEI transfection reagent (Polyplus-transfection, Illkirch, France). Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (both from Biological Industries, Beit Ha'emek, Israel) and maintained at 37 °C and 5% CO_2. Twenty-four hours following transfection, total RNA was extracted from cells with TRI Reagent (Sigma-Aldrich, St. Louis, MO) and treated with RQ1 RNase-free DNase (Promega, Madison, WI). Reverse transcription was performed with 1 µg of DNase-treated total RNA in a 20 µl reaction volume using 200 U of M-MLV Reverse Transcriptase and 100 ng of random primers (Stratagene, La Jolla, CA). Two microliters of cDNA were subjected to PCR amplification with primers located in exons 15 and 16. PCR was performed in a 25 µl reaction volume in the presence of 5X Readymix (LAROVA GmbH, Teltow, Germany) and 10 pmol of each forward and reverse primers. Primer sequences were detailed in the previous paragraph. PCR products were cloned into the pGEM-Teasy vector (Promega). COS-7 cells used for this experiment were authenticated by genotyping 4 STR markers (*D17S1304, D5S1467, D4S2408, D19S245*). Genotyping was performed by PCR amplification of each STR and direct sequencing. PCR was performed on 50 ng of genomic DNA in a 25 µl reaction volume in the presence of 5X Readymix (LAROVA GmbH, Teltow, Germany) and 10 pmol of each forward and reverse primers. PCR primer sequences are provided in Almeida et al [14]. Sequencing was performed with the Big Dye terminator cycle sequencing kit

on an ABI 3130xl Genetic Analyzer (PE Applied Biosystems, Foster City, CA). Results (in terms of repeat number for each STR) were compatible with those described for COS-7 cells [14] (Appendix 1).

Linkage disequilibrium (LD) analysis: DNA samples were from five unrelated patients homozygous for the c.1921–9C>G mutation and from 45 unrelated healthy Jewish subjects of Eastern European origin (Caucasus and Bukhara), who served as controls. All control subjects reported Caucasus Jewish or Bukhara Jewish ancestry on both parental sides. Individuals of mixed ancestry were not included. Five highly polymorphic (expected heterozygosity: 0.63–0.79) dinucleotide short tandem repeats (*STRP: D4S3360/4PTEL04, D4S3038/ AFMa060ze5, D4S43; D4S1614/AFM331xh5*, and *D4S127*) selected through the GeneLoc integrated map of chromosome 4 (v.4.8; Weizmann Institute of Science) were genotyped by PCR amplification with fluorescent dye-labeled primers and capillary electrophoresis (ABI Prism 310; PE Applied Biosystems, Foster City, CA). PCR was performed on 50 ng of genomic DNA in a 15 µl reaction volume in the presence of True Allele PCR Premix (PE Applied Biosystems), and 5 pmol of each forward and reverse primers. Marker haplotypes encompassing the *PDE6B* locus were inferred either by homozygosity or with the use of the expectation-maximization algorithm as implemented in Haploview (v.4.0; Broad Institute, Cambridge, MA) [15]. LD was calculated according to Bengtsson and Thomson's method [16]: $\delta = (p_0 - p_N)/(1 - p_N)$, where p_0 represents the frequency of the allele among individuals with the PDE6B mutation, and p_N represents the frequency of the allele in the general population. Physical map distances were obtained from *GeneLoc* v.4.8 (based on NCBI GRCh38.p12 built), and genetic distances were calculated through a conversion factor derived from the Marshfield genetic map of human chromosome 4 [17]. Mutation age was first estimated with two moment methods [18,19], both of which are based on the "genetic clock." Its equation, $\ln Q = -\theta g$, relates the time (in generations [g]) since the most recent common ancestor (MRCA) of mutant chromosomes, the frequency of recombination between the disease locus and the marker (θ), and the probability that a marker's allele on a disease chromosome is the ancestral one (Q). An unbiased estimate of Q is the proportion of observed haplotypes that are ancestral. The genetic clock was set following Luria and Delbrück's approach [20] which takes into account the population growth rate (d). Accordingly, the estimated age was corrected by adding $g_0 = -(1/d)\ln(\theta/d)$ to the g value [21]. With the moment method introduced by Reich and Goldstein [18], it was possible to correct the decay of LD over generations for the mutation rate ($\mu = 0.00056$ for dinucleotide repeats [22]) at marker loci. LD data were also used to estimate the

MRCA age with Rannala and Reeve's Bayesian Markov chain–Monte Carlo method [23] implemented in the DMLE+ program, v.2.3.

RESULTS

Clinical findings in the studied cohort of patients with RP: A male patient from a non-consanguineous Jewish family who originated from Caucasia was diagnosed with RP at the age of 40 years (patient A-1). He had had night blindness since childhood. At the age of 39 years, his scotopic ERG was non-recordable, and his photopic ERG was significantly reduced. At this age, his BCVA was compromised, and fundus examination revealed pigmentation at the macular area in both eyes (Appendix 2). To further study the genetic basis for RP in Caucasus Jews, we ascertained additional patients with RP of the same ethnic background (Figure 1A).

The diagnosis of RP was primarily based on ERG testing in 6/8 patients (A-1, C-1, C-2, D-1, E-1, and F-1), and was further supported by funduscopic findings and/or visual field testing. In two additional patients (B-1 and E-2), ERG was not performed, and the RP diagnosis was based on fundus findings, visual field constriction, and self-report of night blindness as the initial symptom. All patients had classic symptoms of RP, including night blindness as the initial symptom, concentric restriction of the visual field, and a rod>cone disease process, as indicated by the ERG recordings. Typical funduscopic findings (including narrowed retinal blood vessels and bone spicule-like pigmentation in the periphery) were observed in six patients (A-1, B-1, D-1, E-1, E-2, and F-1). For two patients (C-1 and C-2), funduscopic exam records were not available. Interestingly, significant macular involvement (including macular atrophy, pigmentation at the macular region, and cystic macular edema) was noted in most patients (Appendix 2, Appendix 3, and Figure 2). Age of onset (as reported subjectively by the patients) varied between childhood and 49 years. However, in most patients (6/8), onset was during the fourth decade or later, indicating a relatively late onset of the disease.

Identification of an intronic mutation in PDE6B in a patient with RP with whole exome sequencing: To identify the genetic cause of disease in patient A-1, WES was performed. Because patient A-1 is an isolated case, AR and XL patterns of inheritance were considered. WES resulted in 474,641 variants. Data were initially filtered for rare variants (minor allele frequency ≤1% in the (MAF ≤1% in the 1000 genomes database, the Exome Variant Server and dbSNP) leading to 1,580 missense, nonsense, deletions/insertions, or candidate splicing mutations (based on being located up to 10 bp from an intron–exon junction). We first focused
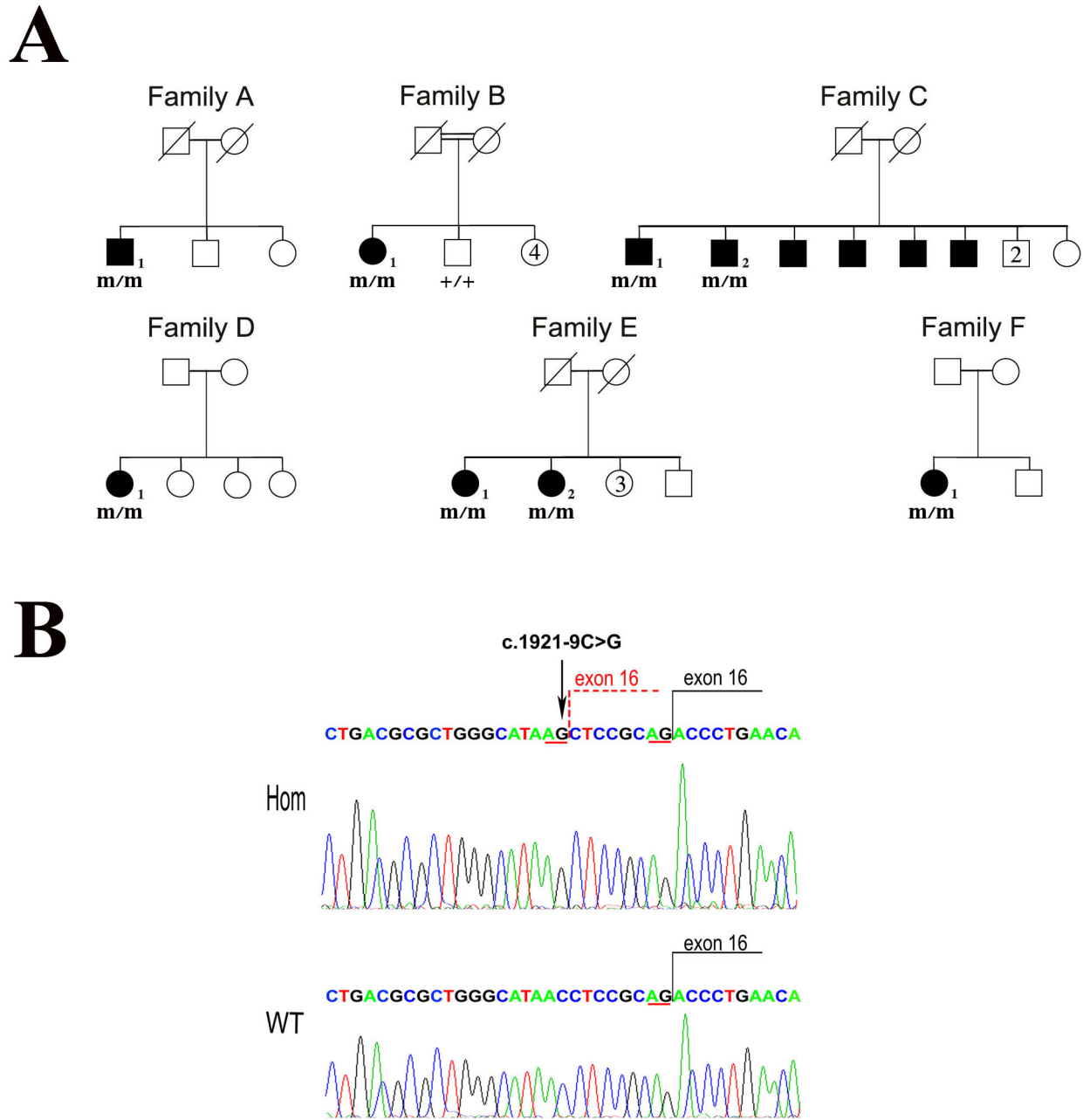
Figure 1. Pedigrees and mutation analysis. **A**: Shown are six unrelated Caucasus Jewish families in whom the c.1921–9C>G mutation is segregated. Filled symbols represent affected individuals, whereas clear symbols represent unaffected individuals. A double line represents consanguinity. Patients recruited for this study are marked by numbers. +, wild-type (wt) allele; m, mutant allele. **B**: Nucleotide sequence traces of the boundary between *PDE6B* intron 15 and exon 16 in a non-carrier individual (wt) and an affected individual homozygote for the c.1921–9C>G mutant allele (hom). The intron–exon boundary is marked. The acceptor splice site (AG) is underlined.

on variants located in known IRD-related genes (RetNet - Retinal Information Network). No putative pathogenic variants (missense, nonsense, deletions/insertions, or canonical (GU-AG) splice-site alterations) in known IRD genes were identified. However, we noted a homozygous C>G transversion, located at position −9 of *PDE6B* intron 15 (NM_000283: c.1921–9C>G; IVS15–9C>G; Figure 1B). This novel variant is not present in the 1000 Genomes database, the Exome Variant Server, dbSNP, or the Genome Aggregation Database (gnomAD).

c.1921–9C>G mutation affects splicing: In silico analysis predicted that c.1921–9C>G generates a novel acceptor splice site, 9 bp upstream from the original splice site of intron 15. *PDE6B* is endogenously expressed mainly in the human retina (and specifically in rod photoreceptors), and to a much lesser extent in other tissues, such as brain, heart, spleen, intestine, prostate, ovary, kidney, and muscle. As these human tissues are not easily accessible, we could not evaluate the effect of c.1921–9C>G on splicing in patient-derived RNA. Alternatively, we used an in vitro splicing assay approach. For this purpose, we created two minigene constructs (wt and mutant) harboring *PDE6B* exons 15 and 16, and the intron between them, downstream of a CMV promoter (Figure 3A). For transfection, we needed a human or primate cell line (which provides the mammalian splicing machinery), which does not express *PDE6B* endogenously (so we can easily detect transcripts obtained from our minigene construct, without an endogenic background). We chose the COS-7 cell line, which is derived from the monkey kidney. Constructs were transfected into COS-7 cells, followed by

RNA extraction and reverse transcription (RT)–PCR with primers located in exons 15 and 16. Both constructs yielded similar products, although the mutant product was slightly bigger (Figure 3B). Both products were subcloned, and eight independent clones of each product were sequenced. Sequencing results revealed that cDNA derived from the wt construct yielded the expected product of 172 bp, in which intron 15 was removed and exons 15 and 16 were correctly spliced. In contrast, cDNA derived from the mutant construct yielded an aberrantly spliced product of 180 bp, in which the newly created acceptor site was used instead of the wt site, leading to an 8-bp insertion into exon 16 (Figure 3A). This insertion is predicted to generate a frameshift and premature termination of translation of the *PDE6B* protein product (p.T641Lfs*5).

Identification of c.1921–9C>G in additional patients: We used PCR amplification and direct sequencing to test 16 additional patients with RP from 14 unrelated Caucasus Jewish families. c.1921–9C>G was identified homozygously in seven additional patients from five families. Only one
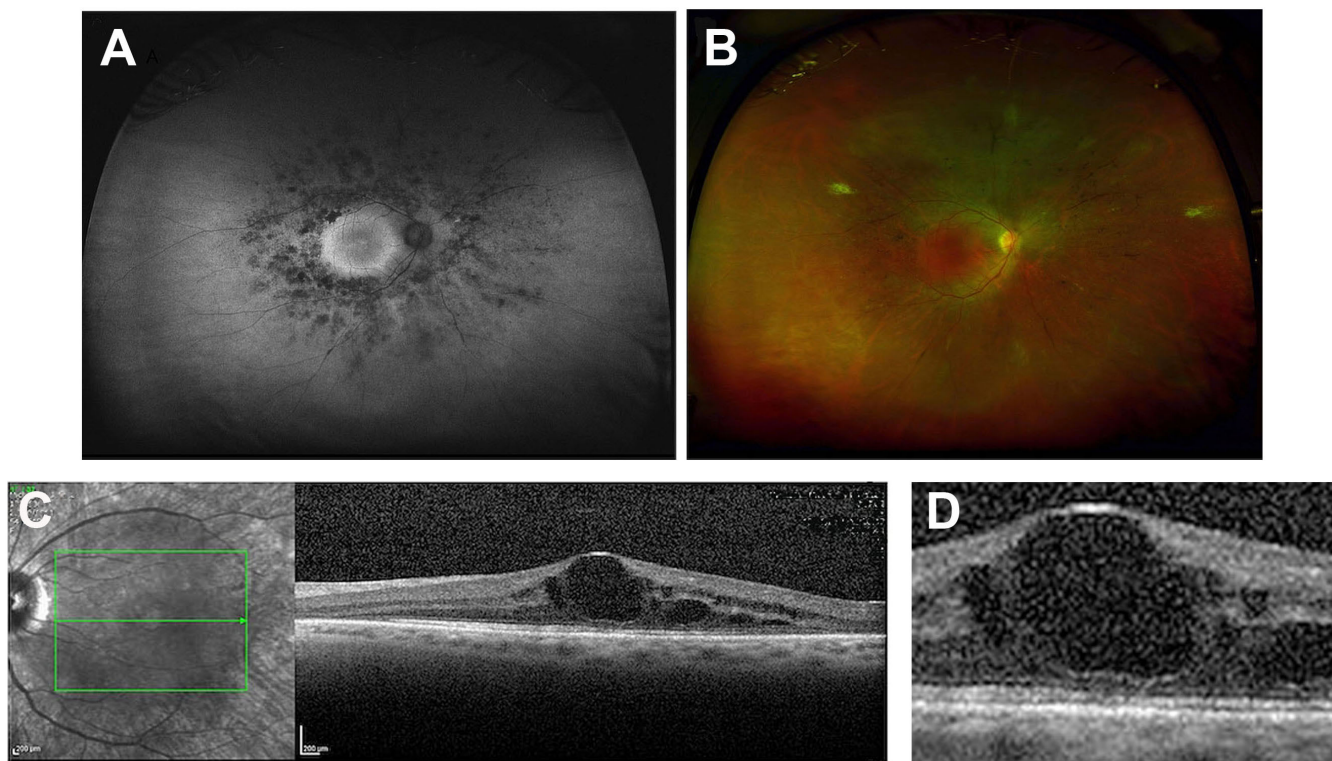

Figure 2. Widefield autofluorescence and pseudocolor fundus imaging and SD–OCT of patient E-2. **A, B**: Autofluorescence and fundus Optos imaging at the age of 57 years, showing an atrophic macula, pigment deposits at the periphery, and attenuated veins. **C**: Autofluorescence and optical coherence tomography (OCT) at the age of 54 years, showing central subfluid thickness on Heidelberg spectral OCT due to severe edema causing elimination of the ellipsoid zone (EZ) temporally and at the foveal center in both eyes. **D**: An inset picture from panel C, showing the magnified fovea and the lost EZ.

of the families was consanguineous (Figure 1A). In total, c.1921–9C>G was identified as the cause of arRP in 40% (6/15) of the Caucasus Jewish families tested.

To test the possibility that the earlier age of onset reported in patients A-1 (childhood) and C-2 (18 years) is due to additional genetic factors, we reviewed the next-generation sequencing results of both patients for additional variants in IRD-related genes. In patient A-1, we observed a heterozygous variant in the *ROM1* gene (OMIM 180721; NM_000327, c. 671C>T, p.Pro224Leu). This rare variant (gnomAD AF=0.00004) is predicted to be pathogenic by several prediction tools. Interestingly, *ROM1* is involved in digenic inheritance of RP, together with *PRPH2* (OMIM 179605; NM_000322), and acts as a modifier of the *PRPH2*-induced retinal phenotype [24,25]. Whether *ROM1* acts as a modifier of *PDE6B* as well is still to be determined.

Age of the founder mutation: The occurrence of the same biallelic *PDE6B* c.1921–9C>G mutation in eight patients with arRP belonging to six independent Caucasus Jewish families is suggestive of a common ancestor who introduced the mutation in the population by a founder effect. We addressed this intriguing question by performing haplotype analysis with the use of five STRPs and the c.1921–9C>G mutation. The presence of an ancestral six-marker haplotype (5-G-8-4-3-5) and several related haplotypes derived from recombination events in the mutation-bearing chromosomes is compatible with the hypothesis that the pathogenic variant c.1921–9C>G can be traced to an ancient common founder belonging to the Caucasus Jewish population ancestry. Thereafter, we estimated the age of the MRCA of the *PDE6B* c.1921–9C>G chromosomes. The presence of significant LD ($\delta > 0.1$) at five highly polymorphic STRP markers (Table 1) allowed the estimation of the coalescence time with two moment methods [18,19], both of which are based on the "genetic clock" [21]. Due to the uncertainty in inferring the historical growth rate of the Caucasus Jewish population, the Luria-Delbrück
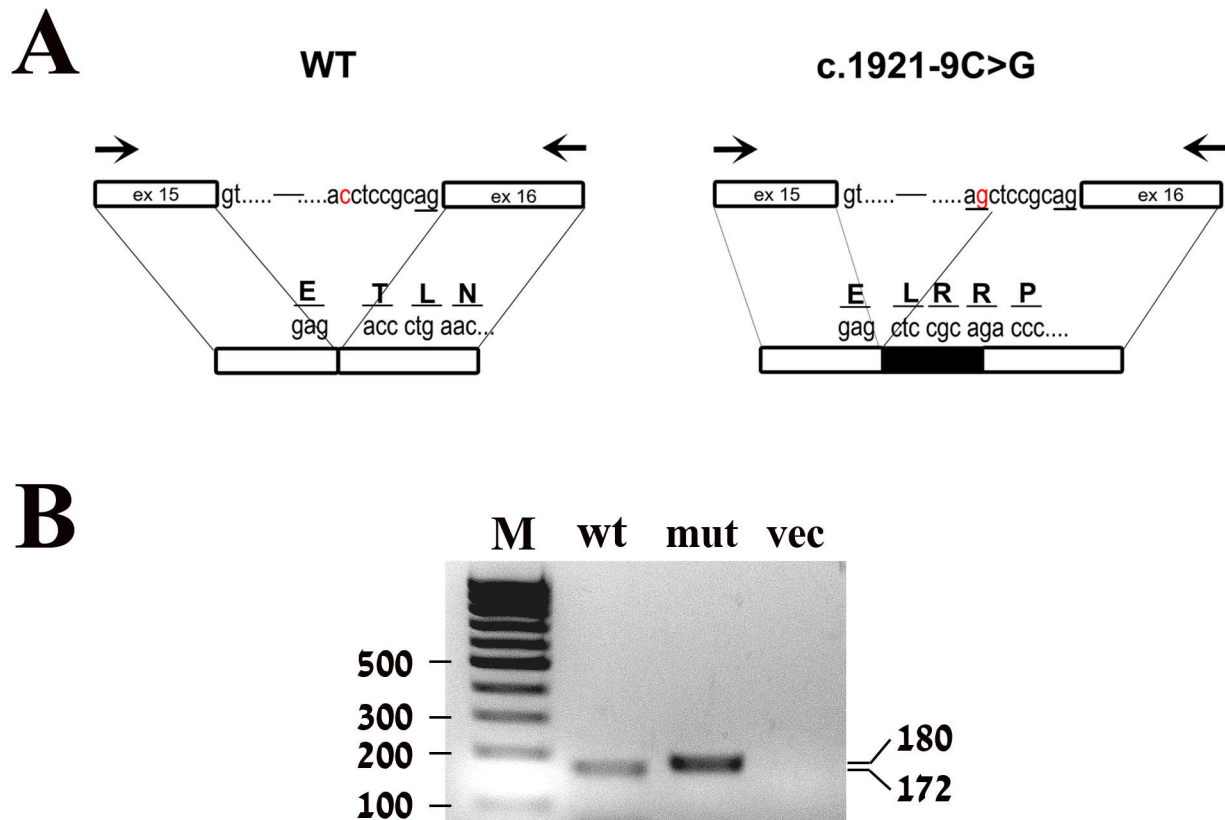

Figure 3. Minigene constructs and products obtained in the in vitro splicing assay. **A**: Shown is a schematic representation of the constructs, which include *PDE6B* exons 15 and 16 (represented by boxes) and the intron between them (represented by a straight line). Either a C or a G is present at position −9 of intron 15. Also shown are the locations of the primers used for reverse transcription (RT)-PCR (indicated by arrows). **B**: Constructs were transfected into COS-7 cells, followed by RNA extraction and RT–PCR. cDNA derived from the wild-type (wt) construct yielded a correctly-spliced product. cDNA derived from the mutant construct yielded an aberrantly spliced product, with an 8-bp insertion into exon 16.

TABLE 1. ESTIMATION BY MOMENT METHODS OF THE AGE OF THE c.1921–9C>G MUTATION IN THE *PDE6B* GENE CAUSING AUTOSOMAL RECESSIVE RETINITIS PIGMENTOSA AMONG JEWS FROM THE CAUCASUS.

Marker	Distance from *PDE6B*			Allele	LD [b]	Population growth rate	Estimated age [a]				
								Method 1		Method 2	
	Mb	cM	θ^{δ}		δ	d	g_0	g_1	g_{1c}	g_2	g_{2c}
D4S3360	0.547	1.69	0.017	5	0.325	0.2	12.3	65.9	78.2	58	70.3
						0.4	7.91	65.9	73.8	58	65.9
D4S3038	0.443	1.37	0.014	8	0.284	0.2	13.4	91.2	104.6	84	97.4
						0.4	8.43	91.2	99.6	84	92.4
D4S43	1.578	4.89	0.049	4	0.188	0.2	7.04	33.3	40.3	31	38
						0.4	5.25	33.3	38.5	31	36
D4S1614	1.982	6.14	0.061	3	0.118	0.2	5.9	33.8	39.7	32	37.9
						0.4	4.68	33.8	38.5	32	36.7
D4S127	2.352	7.29	0.073	5	0.115	0.2	5.05	28.6	33.6	28	33
						0.4	4.26	28.6	32.9	28	32

[a] g_1 is the age (number of generations) calculated by use of the method of Risch et al. [19]; g_2 is the age obtained according to the iterative procedure of Reich and Goldstein [18]. The ages adjusted for population growth rate (g_{1c} and g_{2c}) were obtained by adding the Luria–Delbrück correction factor (g_0) to g_1 and g_{2c}, respectively; g_0 was calculated under the assumptions d=0.2 or=0.4. [b] Linkage disequilibrium, calculated according to Bengtsson and Thomson [16]: $\delta = (p_D - p_N)/(1 - p_N)$. [d] Estimated on the basis of physical distance (Mb) by use of the conversion factor 3.1 cM/Mb (derived from the Marshfield genetic map of human chromosome 4) and Kosambi's map function.

adjustment of the age [20] was applied twice, under the assumption of the extremes of a reasonable interval of the *d* value (0.2–0.4).

The mean (± standard deviation [SD]) overall age estimate for the *PDE6B* c.1921–9C>G mutation is 58.4±27.1 *g* (95% confidence interval [CI]: 39.1–77.8 *g*) with use of the first algorithm [19] and is 53.9±24.1 *g* (95% CI: 36.7–71.2 *g*) with use of the second method [18]. Because simple parametric age estimators based on the genetic-clock equation suffer from the uncertainty about the intra-allelic genealogy, the LD data were reanalyzed from a Bayesian perspective. Applying a mean population growth rate of 0.2–0.4, a Markov chain–Monte Carlo algorithm [23] provided an estimation of 57.3 *g* (95% CI: 45.4–76.2 *g*), an age similar to that obtained by the moment methods. However, the lower limit of the CI is higher than the corresponding confidence limits provided by these methods.

DISCUSSION

The aim of the present study was to identify the genetic basis for arRP in a cohort of Caucasus Jewish patients. Genetic analysis revealed a novel intronic mutation in *PDE6B*, c.1921–9C>G, leading to altered splicing. The in vitro splicing assay we performed confirmed that the c.1921–9C>G variant creates a novel acceptor splice-site, which is efficiently recognized by the primate splicing machinery. However, the exact effect of this splicing mutation on *PDE6B* transcripts in the human retina is not known, and may differ from the effect we observed in COS-7 cells, due to tissue-specific differences in the splicing machinery.

Based on the present findings, c.1921–9C>G is a major cause of arRP among Caucasus Jews, as this mutation accounts for the disease in up to 40% of the families. c.1921–9C>G is a rare mutation, which is not present in public databases, and was detected only in patients of Caucasus Jewish ancestry. All patients with the c.1921–9C>G mutation had classic symptoms of RP, with onset that varied between childhood and 49 years. Clinical findings in these patients were similar to those reported previously in patients with other *PDE6B* mutations [2,3,26,27].

We used ANNOVAR software [28] as part of our customized pipeline. ANNOVAR is an extremely popular software tool (cited more than 3,000 times). The default threshold in ANNOVAR for variant distance from the intron–exon junction is 2 bp. In the analysis, we increased this threshold to 10 bp, which allowed the identification of the c.1921–9C>G mutation. The conclusion is that when ANNOVAR is used with the default threshold, many mutations located proximate

to the intron–exon junction, but not directly within the canonical (GU-AG) splice-sites, are missed. In this context, currently, despite the introduction of next-generation sequencing techniques, reported detection rates for large IRD cohorts reach approximately 70% [29-31]. Many of the missing mutations probably reside within introns. Some intronic mutations, not residing within canonical splice sites but affecting splicing, are common causes of IRD in certain populations (e.g., the c.7595–2144A>G mutation in *USH2A* (OMIM 608400; NM_206933) and the c.4539+2001G>A and c.4539+2028C>T mutations in *ABCA4;* OMIM 601691; NM_000350) [32,33].

Homozygosity for the c.1921–9C>G mutation among Caucasus Jewish patients with RP is usually not due to consanguinity (as indicated by the fact that only one of the six positive families was consanguineous). Instead, homozygosity reflects a founder effect, as demonstrated by the identification of an ancestral haplotype spanning 9 cM and encompassing the *PDE6B* c.1921–9C>G mutation on chromosome 4p16.3 in these patients. The decay of LD over generations, modeled through different algorithms, would date the age of the MRCA of the mutation-bearing chromosomes to 1,300–1,500 years ago, assuming 25 years per generation. Throughout the calculations, we assumed that the generations were non-overlapping, there is no population substructure, the mutant allele is selectively neutral, and there is no further recurrent mutation at the disease locus. According to Kaplan et al. [34] and Guo and Xiong [35], in the absence of any evidence to the contrary, it is reasonable to assume that a mutation causing a rare AR disorder is selectively neutral. A population growth rate of 0.2–0.4 was applied as a reasonable interval based on available historical Jewish census data.

Historians have offered several hypotheses about the origins of the first Mountain Jewish communities in the Caucasus (reviewed in [9]). According to one hypothesis, the first Caucasian Jewish community was made up of refugees from the southwestern area of Sassanian Iran and dates back to the 6th century CE. Assuming 25 years per generation, the mean of the estimated age of the MRCA of the c.1921–9C>G mutation-bearing chromosomes (56 *g*) points to the middle of that century as the most probable date of a founder effect, the pathogenic *PDE6B* variant being introduced into the expanding settlement by a displaced Jewish family. Considering the 95% confidence interval generated by the Bayesian approach to the age inference, the lower value (45 *g*; 1,125 years) is consistent with a founder effect originating in the 9th century CE, when many Jews from the Byzantine Empire moved to Khazaria and eventually, the Caucasus because the

Christian Emperor Justinian persecuted them. At the opposite end of the confidence interval (76 *g*; 1,900 years), MRCA age generates a date that overlaps with the time of the Roman destruction of the Temple and the beginning of the second Jewish diaspora (1st century CE), which may have created an earlier occasion for the introduction of the c.1921–9C>G mutation into the Caucasus Jewish population.

The molecular causes of RP are strikingly heterogeneous, with more than 60 genes associated with arRP. The contribution of each of these genes to the overall prevalence of disease is relatively small, and for many of them, pathogenic mutations have been reported in only a few families worldwide. Due to this situation, genetic analysis in patients with RP and their families is challenging. Identification of founder mutations that are common in specific ethnic groups allows for rapid, sensitive, specific, and cost-effective use of genetic testing for carrier screening and diagnostic purposes. For example, some of the common founder mutations we identified in certain Jewish subpopulations are routinely tested by Israeli genetic institutes [36,37]. Moreover, until recently there was no effective treatment for RP. However, the field has undergone dramatic changes in the last decade, including development of therapeutic modalities that are based on the available knowledge of the causative gene (e.g., gene therapy) [38]. Therefore, identification of the underlying gene in a patient with RP may facilitate treatment as well. A clinical trial testing the safety and efficacy of gene therapy in patients with RP due to *PDE6B* mutations is being conducted in France (NCT03328130, https://clinicaltrials.gov).

In summary, the identification of the c.1921–9C>G mutation of *PDE6B* as a major cause of arRP in Caucasus Jews will facilitate molecular diagnosis, carrier screening, and genetic counseling, as well as treatment, in this population.

APPENDIX 1. AUTHENTICATION OF COS-7 CELLS.

To access the data, click or select the words "Appendix 1."

APPENDIX 2. CLINICAL FINDINGS IN PATIENTS HOMOZYGOTES FOR THE C.1921-9C>G MUTATION

To access the data, click or select the words "Appendix 2."

APPENDIX 3. ERG RESPONSES.

To access the data, click or select the words "Appendix 3."

REFERENCES

1. Verbakel SK, van Huet RAC, Boon CJF, den Hollander AI, Collin RWJ, Klaver CCW, Hoyng CB, Roepman R, Klevering BJ. Non-syndromic retinitis pigmentosa. Prog Retin Eye Res 2018; •••:[PMID: 29597005].

2. Danciger M, Blaney J, Gao YQ, Zhao DY, Heckenlively JR, Jacobson SG, Farber DB. Mutations in the PDE6B gene in autosomal recessive retinitis pigmentosa. Genomics 1995; 30:1-7. [PMID: 8595886].

3. McLaughlin ME, Sandberg MA, Berson EL, Dryja TP. Recessive mutations in the gene encoding the beta-subunit of rod phosphodiesterase in patients with retinitis pigmentosa. Nat Genet 1993; 4:130-4. [PMID: 8394174].

4. Ali S, Riazuddin SA, Shahzadi A, Nasir IA, Khan SN, Husnain T, Akram J, Sieving PA, Hejtmancik JF, Riazuddin S. Mutations in the beta-subunit of rod phosphodiesterase identified in consanguineous Pakistani families with autosomal recessive retinitis pigmentosa. Mol Vis 2011; 17:1373-80. [PMID: 21655355].

5. Carss KJ, Arno G, Erwood M, Stephens J, Sanchis-Juan A, Hull S, Megy K, Grozeva D, Dewhurst E, Malka S, Plagnol V, Penkett C, Stirrups K, Rizzo R, Wright G, Josifova D, Bitner-Glindzicz M, Scott RH, Clement E, Allen L, Armstrong R, Brady AF, Carmichael J, Chitre M, Henderson RH, Hurst J, MacLaren RE, Murphy E, Paterson J, Rosser E, Thompson DA, Wakeling E, Ouwehand WH, Michaelides M, Moore AT. Consortium NI-BRD, Webster AR, Raymond FL. Comprehensive Rare Variant Analysis via Whole-Genome Sequencing to Determine the Molecular Pathology of Inherited Retinal Disease. Am J Hum Genet 2017; 100:75-90. [PMID: 28041643].

6. Kim C, Kim KJ, Bok J, Lee EJ, Kim DJ, Oh JH, Park SP, Shin JY, Lee JY, Yu HG. Microarray-based mutation detection and phenotypic characterization in Korean patients with retinitis pigmentosa. Mol Vis 2012; 18:2398-410. [PMID: 23049240].

7. Gal A, Orth U, Baehr W, Schwinger E, Rosenberg T. Heterozygous missense mutation in the rod cGMP phosphodiesterase beta-subunit gene in autosomal dominant stationary night blindness. Nat Genet 1994; 7:64-8. [PMID: 8075643].

8. Manes G, Cheguru P, Majumder A, Bocquet B, Senechal A, Artemyev NO, Hamel CP, Brabet P. A truncated form of rod photoreceptor PDE6 beta-subunit causes autosomal dominant congenital stationary night blindness by interfering with the inhibitory activity of the gamma-subunit. PLoS One 2014; 9:e95768-[PMID: 24760071].

9. Farzaliev S. Judaism in the Caucasus. Historical Essay. The Caucasus & Globalization 2008; 2:148-57. .

10. Robson AG, Nilsson J, Li S, Jalali S, Fulton AB, Tormene AP, Holder GE, Brodie SE. ISCEV guide to visual electrodiagnostic procedures. Doc Ophthalmol 2018; 136:1-26. [PMID: 29397523].

11. Grimberg J, Nawoschik S, Belluscio L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acids Res 1989; 17:8390-[PMID: 2813076].

12. Eytan O, Morice-Picard F, Sarig O, Ezzedine K, Isakov O, Li Q, Ishida-Yamamoto A, Shomron N, Goldsmith T, Fuchs-Telem D, Adir N, Uitto J, Orlow SJ, Taieb A, Sprecher E. Cole Disease Results from Mutations in ENPP1. Am J Hum Genet 2013; 93:752-7. [PMID: 24075184].

13. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 2009; 37:e67-[PMID: 19339519].

14. Almeida JL, Hill CR, Cole KD. Authentication of African green monkey cell lines using human short tandem repeat markers. BMC Biotechnol 2011; 11:102-[PMID: 22059503].

15. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21:263-5. [PMID: 15297300].

16. Bengtsson BO, Thomson G. Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens 1981; 18:356-63. [PMID: 7344182].

17. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 1998; 63:861-9. [PMID: 9718341].

18. Reich DE, Goldstein DB. Estimating the age of mutations using variation at linked markers. In: Goldstein, GB,Schlötterer, C, editors. Microsatellites: evolution and applications. UK: Oxford University Press; 1999. p.129–38.

19. Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, Breakefield X, Bressman S. Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat Genet 1995; 9:152-9. [PMID: 7719342].

20. Luria SE, Delbruck M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics 1943; 28:491-511. [PMID: 17247100].

21. Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. Am J Hum Genet 1996; 59:633-43. [PMID: 8751865].

22. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet 1993; 2:1123-8. [PMID: 8401493].

23. Rannala B, Reeve JP. High-resolution multipoint linkage-disequilibrium mapping in the context of a human genome

sequence. Am J Hum Genet 2001; 69:159-78. [PMID: 11410841].

24. Conley SM, Stuck MW, Watson JN, Naash MI. Rom1 converts Y141C-Prph2-associated pattern dystrophy to retinitis pigmentosa. Hum Mol Genet 2017; 26:509-18. [PMID: 28053051].

25. Kajiwara K, Berson EL, Dryja TP. Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. Science 1994; 264:1604-8. [PMID: 8202715].

26. Danciger M, Heilbron V, Gao YQ, Zhao DY, Jacobson SG, Farber DB. A homozygous PDE6B mutation in a family with autosomal recessive retinitis pigmentosa. Mol Vis 1996; 2:10-[PMID: 9238087].

27. Tsang SH, Tsui I, Chou CL, Zernant J, Haamer E, Iranmanesh R, Tosi J, Allikmets R. A novel mutation and phenotypes in phosphodiesterase 6 deficiency. Am J Ophthalmol 2008; 146:780-8. [PMID: 18723146].

28. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38:e164-[PMID: 20601685].

29. Riera M, Navarro R, Ruiz-Nogales S, Mendez P, Bures-Jelstrup A, Corcostegui B, Pomares E. Whole exome sequencing using Ion Proton system enables reliable genetic diagnosis of inherited retinal dystrophies. Sci Rep 2017; 7:42078-[PMID: 28181551].

30. Stone EM, Andorf JL, Whitmore SS, DeLuca AP, Giacalone JC, Streb LM, Braun TA, Mullins RF, Scheetz TE, Sheffield VC, Tucker BA. Clinically Focused Molecular Investigation of 1000 Consecutive Families with Inherited Retinal Disease. Ophthalmology 2017; 124:1314-31. [PMID: 28559085].

31. Tiwari A, Bahr A, Bahr L, Fleischhauer J, Zinkernagel MS, Winkler N, Barthelmes D, Berger L, Gerth-Kahlert C, Neidhardt J, Berger W. Next generation sequencing based identification of disease-associated mutations in Swiss

patients with retinal dystrophies. Sci Rep 2016; 6:28755-[PMID: 27353947].

32. Albert S, Garanto A, Sangermano R, Khan M, Bax NM, Hoyng CB, Zernant J, Lee W, Allikmets R, Collin RWJ, Cremers FPM. Identification and Rescue of Splice Defects Caused by Two Neighboring Deep-Intronic ABCA4 Mutations Underlying Stargardt Disease. Am J Hum Genet 2018; 102:517-27. [PMID: 29526278].

33. Slijkerman RW, Vache C, Dona M, Garcia-Garcia G, Claustres M, Hetterschijt L, Peters TA, Hartel BP, Pennings RJ, Millan JM, Aller E, Garanto A, Collin RW, Kremer H, Roux AF, Van Wijk E. Antisense Oligonucleotide-based Splice Correction for USH2A-associated Retinal Degeneration Caused by a Frequent Deep-intronic Mutation. Mol Ther Nucleic Acids 2016; 5:e381-[PMID: 27802265].

34. Kaplan NL, Hill WG, Weir BS. Likelihood methods for locating disease genes in nonequilibrium populations. Am J Hum Genet 1995; 56:18-32. [PMID: 7825575].

35. Guo SW, Xiong M. Estimating the age of mutant disease alleles based on linkage disequilibrium. Hum Hered 1997; 47:315-37. [PMID: 9391824].

36. Auslender N, Bandah D, Rizel L, Behar DM, Shohat M, Banin E, Allon-Shalev S, Sharony R, Sharon D, Ben-Yosef T. Four USH2A founder mutations underlie the majority of Usher syndrome type 2 cases among non-Ashkenazi Jews. Genet Test 2008; 12:289-94. [PMID: 18452394].

37. Auslender N, Sharon D, Abbasi AH, Garzozi HJ, Banin E, Ben-Yosef T. A common founder mutation of CERKL underlies autosomal recessive retinal degeneration with early macular involvement among Yemenite Jews. Invest Ophthalmol Vis Sci 2007; 48:5431-8. [PMID: 18055789].

38. Lee JH, Wang JH, Chen J, Li F, Edwards TL, Hewitt AW, Liu GS. Gene therapy for visual loss: Opportunities and concerns. Prog Retin Eye Res 2019; 68:31-53:[PMID: 30170104].