Contribution of structural and intronic mutations to *RPGRIP1*-mediated inherited retinal dystrophies.

Farzad Jamshidi[1], Emily M. Place[1], Daniel Navarro-Gomez[1], Mathew Maher[1], Elise Valkanas[3], Monkol Lek[2,3], Daniel MacArthur[2,3], Kinga M. Bujakowska[1], Eric A. Pierce[1]

[1]Ocular Genomics Institute, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA.

[2]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA.

[3]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

Correspondence:

Eric A. Pierce M.D., Ph.D.

Solman and Libe Friedman Professor of Ophthalmology
Massachusetts Eye and Ear Infirmary
Harvard Medical School
243 Charles Street
Boston, MA 02114
Phone: 617-573-6917
eric_pierce@meei.harvard.edu

1. Abstract

With the completion of the first phase 3 human gene therapy randomized clinical trial, in the form of voretigene neparvovec for *RPE65*-mediated inherited retinal dystrophy, as well as the advent of more than 10 other gene therapy trials for inherited retinal disorders, accurate genetic diagnostics will have an increasingly important role in clinical decision-making. Current genetic diagnostic testing panels primarily focus on coding sequences. However, we find that structural and intronic variations are crucial in solving ambiguous cases. We present four families in whom more in depth sequencing and bioinformatic analyses led to the identification of non-coding and structural variations in *RPGRIP1* as the cause of disease. In the process we describe ten novel *RPGRIP1* mutations. We suggest an expansion of both sequencing and bioinformatics tools in the diagnostics of inherited retinal degenerations to increase sensitivity and to make confident calls before enrolling patients in specific gene therapy clinical trials.

2. Introduction

In 1869, Theodor Leber while consulting for the Ilvesheim School for the Blind described a type of blindness occurring at birth or soon after accompanied with nystagmus, unexpected photophobia, and a sluggish pupillary response (1). Diminished electro-retinogram (ERG) signals, along with an autosomal recessive pattern of inheritance were associated with similar cases in the 1950s (2-4) and this group of early-onset inherited retinal degenerations (IRDs) came to be classified as Leber congential amaurosis (LCA). It has a prevalence of 1 in 33000 to 81000, and comprises an estimated 18% of blind children and 5% of IRDs (2-5). Mutations in 25 genes have been linked to LCA and can explain up to 70% of cases (5-38). It should be noted however that mutations in LCA associated genes can also cause other IRD phenotypes and thus that inclusion of the genetic cause of disease in patient's diagnoses is a desirable goal (5, 39). One of the historically LCA-associated genes is *RPGRIP1*, mutations in which are reported to cause 4-5% of LCA (32, 33) and thus approximately 0.2% of IRDs. The majority of *RPGRIP1* mutations create loss-of-function (LoF) alleles, consistent with the autosomal recessive mode of inheritance of *RPGRIP1*-associated IRD (40).

RPGRIP1 is located on 14q11 (41-43) with a coding sequence composed of 3861 base pairs distributed over 24 exons (33) in its largest transcript variant. It encodes a 1287 amino acid protein that was first discovered via two-hybrid assays through its interaction with the Retinitis Pigmentosa GTPase Regulator (RPGR) protein. The RPGR gene is also an IRD disease gene (41-43). The expression of *RPGRIP1* is limited to the retina and testis (41) and it has species and isoform specific localizations. While it primarily localizes to the transition zones of rods and cones in mice, in human and bovine retinas,

various *RPGRIP1* isoforms can be found in the outer segment, along the microtubules and the cytoskeleton of rods and cones as well as in the amacrine cells of the inner plexiform layer (44, 45). RPGRIP1 has been shown to play a critical role in RPGR tethering in the transition zone, opsin trafficking, outer-segment disc organization and photoreceptor survival (46, 47).

RPGRIP1-associated IRDs are particularly attractive candidates for gene therapy with already established success in murine (48) and canine (49) models of disease. Therefore confident genetic diagnosis with *RPGRIP1* as the causal gene will be crucial for effective gene therapy clinical trials. Here we have used a comprehensive genomic approach, including whole genome sequencing and structural analyses, in families in whom routine targeted screening of the *RPGRIP1* coding region identified only one pathogenic allele and thus the causation was not fully established. We also report six novel *RPGRIP1* mutations in the coding region of the gene.

3. Results

The participants of this study suffered from IRD with typical clinical findings demonstrated in Figure 1 and outlined in Table 1. Out of 1719 families who underwent targeted next generation sequencing (NGS) with the Genetic Eye Disease (GEDi) panel (50), seven families had at least one inactivating mutation in *RPGRIP1* and no additional likely pathogenic mutations in other IRD genes (Table 1, Figure 2). Among these, three families had bilallelic *RPGRIP1* mutations (Table 1). With the expanded approach utilizing whole genome (WGS) and whole exome sequencing (WES), the remaining four families were found to have a second mutant *RPGRIP1* allele in the form of a rare non-coding change or a copy number variation (Table 1, Figure 2).

Copy number variations. In family OGI281, the first allele identified via GEDi testing in the two affected siblings was an out of frame deletion (c.1615_1624del10, p.E539Qfs*2) which creates a LoF allele. Analysis of structural variations from WGS (51) revealed a tandem duplication that includes exon 2 in the two affected and one unaffected sibling, confirming that the two mutations are in *trans* (Figure 3). PCR using primers across the predicted region was carried out which revealed the defective allele in all three members of OGI281 while it did not produce a product in control DNA from HEK293T (Figure 3C). Sanger sequencing across the tandem duplication, identified the break-point further confirming these results (Figure 3D). The duplication was confirmed by quantitative real-time (qRT)PCR (Figure 4A). It is important to note that exon 2 is out of frame and hence the resultant mRNA from the defective allele would be non-functional with a premature stop codon that would lead to non-sense mediated decay (NMD) (52).

In family OGI237, a duplication of exons 1 and 2 was identified in the proband and his mother via WES, which were subsequently qRT-PCR validated (Figure 4B). The father contributed a frameshift allele (c.3793-3794insGAAA, p.V1265GfsTer19) as identified via GEDi and exome sequencing (Figure 4B, Table 1). Exons 1 and 2 together are out of frame and their duplication is expected to lead to a frame shift and NMD of the defective mRNA (52).

In subject D1944, who carried a premature termination variant (c.2302C>T, p.R768*), a deletion of exon 19 was detected through coverage based copy number analysis of GEDi data (50, 53). This CNV was confirmed via qRT-PCR (Figure 4B). Exon 19 is out-of-frame hence the deletion creates LoF allele as well.

I**ntronic mutation leading to splicing defect.** In OGI827, a 2bp out-of-frame insertion in the coding sequence of *RPGRIP1*, inherited from the mother, was identified (c.895-896del2, p.G299Sfs*21) through GEDi sequencing. This mutation was clearly disruptive to gene function. However, a second coding mutation was not identified. Rather, an intronic variant c.2367+23delG was detected by GEDi sequencing (Figure 5A). There were no clear splicing signals affected at this site, and the Human Splicing Finder (54) did not predict a significant change in splicing (Supplementary Figure 2). However, given the low allele frequency of this variant in the population, at 0.00023 as per the GnmoAD database (55), we decided to further evaluate this variant.

Since *RPGRIP1* is not expressed in blood or readily available tissues, a mini-gene approach was adopted to study this mutation. The WT, positive control with an artificially created essential splice site mutation (c.2367+1G>A), as well as the mutant

allele (c.2367+23delG) were transfected separately into HEK293T cells and the spliced products analyzed (Supplementary Figure 3). Multiple transcript variants were expressed from the *RPGRIP1* WT mini-gene yet the most prominent species corresponded to the correct splicing-out of introns 14 and 15 (Figure 5B). The positive control, c.2367+1G>A, had retention of intron 15, which was confirmed by Sanger sequencing (Figure 5C). The c.2367+23delG mutant showed the presence of both the fully spliced and the intron 15 containing transcript variants, with the aberrant transcript being more prominent (Figure 5B). This result confirmed the pathogenic effect of c.2367+23delG leading to aberrant splicing of *RPGRIP1* intron 15.

4. Discussion

Accurate genetic diagnostics for inherited retinal degenerations are becoming more and more critical as they will have increasingly significant therapeutic implications. The results of the first phase 3 human gene therapy randomized clinical trial, voretigene neparvovec for the treatment of LCA2, has been recently published (56). Furthermore, more than 10 clinical trials for retinal degeneration have either been completed or initiated (57-63). A multitude of additional candidates for gene therapy with already established animal models exist and *RPGRIP1* is one such example (48, 49). However, evidence of biallelic gene mutations is critical for inclusion of patients in trials when dealing with such autosomal recessive diseases. The prominent focus of clinical diagnostics has been on analyzing the coding sequence of known IRD genes using panel sequencing (64-68). More recently, the importance of adding copy number analysis has been highlighted and implemented by some investigators (53, 69, 70). In our analysis of 7 *RPGRIP1* mutation families, we find 57% (4/7) of the time, the second mutant allele involves structural variation or intronic mutations that need additional sequencing, bioinformatics analysis and functional validation.

Our finding of 3 families out of 1719 (0.17%) with biallelic inactivation in the coding sequence corresponds with the reported rate of *RPGRIP1*-mutant IRDs in the literature, that is ~0.2% (5, 63, 71). While most of the coding mutations led to frameshifts and premature terminations affecting a significant portion of the encoded protein (Figure 6), the c.3793ins4 (p.V1265Gfs*19) variant occurs in the last exon hence its significance was not as clear. However, given the extremely low allele frequency in the population (0.000004, corresponding to 1 allele in Gnomad database, Table 1) and the resultant

perturbation of the last 21 amino acids of the critical RPGR interacting domain (RID) (72) of *RPGRIP1*, this mutation was deemed significant.

In addition to the three families with two coding mutations in *RPGRIP1*, GEDi testing identified 4 families with single apparent mutations in *RPGRIP1*. We identified and validated novel non-coding *RPGRIP1* mutations in these four families (Figure 6, Table 1). This more than doubled the number of *RPGRIP1*-mediated IRD families initially identified via coding sequence analysis only. In three of these families, copy number changes and out-of-frame deletions or tandem duplications were identified. Others have also reported the importance of adding copy number analysis for IRD diagnostics (53, 69, 70). It is worth noting that one of the structural variants identified in these studies was not detected by exon-based analyses, suggesting that in some cases WGS will be needed to identify mutant alleles (73).

In addition to copy number changes, we identified a single intronic base pair deletion in an *RPGIP1* allele (c.2367+23delG) that was proven *in vitro* to lead to changes in splicing patterns (Figure 5 and Supplementary Figure 4). Our approach to validating this mutation was to generate a minigene assay and use the splicing machinery of HEK293T cells to study the effect of the discovered mutation on splicing patterns. It is important to note that splicing patterns can be distinct depending on tissue type (74) and HEK293T cells are not the target cells in IRD. Yet due to the ease of transfection and rapid growth, they are an excellent *in vitro* tool and can identify significant mutations as evidenced by our positive control essential splice site mutation. Another approach could be the use of RNA from the patient's lymphocytes or other cells such as fibroblasts and examining splicing variants directly from the RNA (75). However, this can prove difficult when dealing with

genes specifically expressed in the retina.

Empirical validation of non-coding variants can be time consuming and expensive, yet the proof of biallelic gene inactivation in recessive disease is essential for provision of gene-specific treatments, such as gene therapies. Statistical models predict a small false positive diagnosis rate when only a single mutation is identified in a disease-specific gene such as *MYO7A* in Usher Syndrome Type I (76). But when a large number of potential genes can lead to the phenotype of interest, such as in LCA/IRD, a larger false positive rate is expected (76). Thus we believe that a thorough characterization of potentially deleterious alleles is crucial for correct genetic diagnosis and useful for increasing the number of patients eligible for a particular genetic therapy.

In short, here we show that with extended sequencing and bioinformatics approaches, a significant portion of inactivating alleles can be identified. In the process we described six novel coding and four novel non-coding mutations in *RPGRIP1*, two of which involve duplications in exon 2 (Figure 6). We focused our research on *RPGRIP1*, after having noticed a relatively high frequency of single loss of function alleles in this gene (Figure 6, Table 1). Our findings point toward a significant proportion of *RPGRIP1*-mediated cases that are missed through routine diagnostic panel sequencing because of more complex structural or intronic changes (Figure 2). This means that the previously estimated prevalence of *RPGRIP1*-mediated IRD is a significant underestimate. Based on our experience, inclusion of more in-depth copy number and intronic analysis can increase the number of potential *RPGRIP1* gene therapy clinical trial participants by a factor of two, which is significant for this rare disease.

5. Methods

Ethical Guidelines. The study was approved by the institutional review board at the Massachusetts Eye and Ear Infirmary (Human Studies Committee MEEI in USA) and adhered to the Declaration of Helsinki. Informed consent was obtained from all individuals on whom genetic testing and further molecular evaluations were performed.

Clinical evaluation. All the patients in this study underwent expert clinical assessment by an ophthalmologist sub-specializing in inherited retinal degenerations. The clinical characteristics are outlined in Table 1.

Sequencing. All samples underwent GEDi sequencing as described previously (50). Simply put, the coding exons, 5′ and 3′-untranslated regions, and select deep intronic regions of more than 238 inherited retinal degeneration genes were captured using SureSelect Target Enrichment (Agilent Technologies, Santa Clara, CA). Samples were indexed, multiplexed and sequenced on a MiSeq platform (Illumina, San Diego, CA). Whole exome and whole genome sequencing were performed at the Broad Institute on HiSeq 4000 NGS instruments (Illumina, San Diego, CA) and HiSeq X (Illumina, San Diego, CA) respectively (http://genomics.broadinstitute.org/). All of the above sequencing was done on DNA extracted from venous blood using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Sanger sequencing was performed on ABI 3730xl using BigDye Terminator v3.1 kits (Life Technologies, Carlsbad, CA) and relevant PCR primers as per Supplementary Table 1. When PCR products were sequenced, they were cleaned to remove primers prior to Sanger sequencing (ExoSap-IT, Affymetrix, Santa Clara, CA). Gel bands that were Sanger sequenced had

DNA extracted via the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA).

Bioinformatics. Analysis was done as described previously (50, 77). BWA was used for alignment. SAMtools and custom programs were used for single nucleotide polymorphism and insertion/deletion calls (77). Variants of interest were limited to polymorphisms with less than 0.005 allelic frequency in the gnomAD (http://gnomad.broadinstitute.org/) and ExAC (http://exac.broadinstitute.org/) databases(55). Whole genome copy number analysis, with consideration of structural changes, was carried done via Genome STRiP 2.0 (51).

Cell culture and transfections. Human embryonic kidney cells (HEK293T) purchased from American Type Culture Collection (ATCC, Manassas, Virginia) and maintained in Dulbecco's Modified Eagle Medium (DMEM): Nutrient Mixture F-12 (Gibco, Live Technologies), supplemented with 10% fetal bovine serum (Gibco). 2ml of 5 x 10^5 cells/ml were plated into each well of a 6-well plate (Corning Inc., Corning, NY) 12hrs prior to transfections. 1 μg of vector DNA combined with 12ul of Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) were combined in Opti-MEM media (Thermo Fisher Scientific, Waltham, MA) and added to each well after 15 minutes of incubation. After 48 h, RNA was extracted and converted to cDNA, which was then used for reverse transcription-polymerase chain reaction (RT-PCR) to analyze splicing patterns.

RNA Isolation and cDNA Syntehsis. Cells were treated with TRIzol (Thermo Fisher Scientific, Waltham, MA). The aqueous phase was then transferred to mRNeasy kit

(Qiagen, Hilden, Germany) and RNA extracted as per protocol. The samples underwent on-column DNase I digestion for a minimum of 15 minutes. Quantification was via NanoDrop (Thermo Fisher Scientific, Waltham, MA) and 500ng of RNA was converted to cDNA using oligo(dT) primers and SuperScript II (Thermo Fisher Scientific, Waltham, MA). Identical reactions without the addition of SuperScript II reverse reverse transcriptase (RT -) were used for control.

Polymerase chain reactions, cloning and site-directed mutagenesis. Exons 11 to 16 of *RPGRIP1*, with at least a 100bp intronic extensions on both the 5' and 3' ends were cloned into an expression vector using OGI281 proband's father's DNA as a template (Supplementary Figure 3). This allowed us to separately clone a WT allele, as well as the mutant (c.2367+23delG) which were separately transfected into HEK293T cells in the pCS2 expression vectors. After 48 hours, the isolated mRNA was converted to cDNA and PCR was performed as shown in Figure 6B. As a control, site-directed mutagenesis was used to introduce a mutation in the essential splice site of the WT at the donor site of exon 15 (c.2367+1G>A). Thermal cycling was done C1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA). For cloning, primers with complementary sequences outside of exons 11 and 16 of *RPGRIP1* were used (Supplementary Table 1) and polymerase chain reaction (PCR) was performed using PfuUltra II Fusion polymerase (Agilent Technologies, Santa Clara, CA). Annealing was done at 60°C with extension performed for 15s/Kb. The PCR products were then cloned into pENTR Directional TOPO vector (Thermo Fisher Scientific, Waltham, MA) and used to transform chemically competent *Escherichia coli* (One Shot TOP10, Fisher Scientific, Waltham, MA). The bacteria were plated on Luria Bertani (LB) agarose plates with Kanamycin selection,

and grown overnight. Up to 15 colonies were selected, DNA extracted via ZymoPURE miniprep kits (Zymo Research, Irvine, CA) and analyzed via Sanger sequencing as well as restriction enzyme digestion with BsrGI (NE Biolabs, Ipswich, MA). Colonies with the correct sequence and restriction enzyme pattern were then sub-cloned into the PSC2 vector via Gateway LR clonase II (Thermo Fisher Scientific, Waltham, MA) and similar analysis as before was done to isolate vector with the appropriate insert for transfection experiments. PCR for evaluative purposes was done using HOT FIREPol^{R} DNA Polymerase (Solis BioDyne, Tartu, Estonia) with annealing at 60°C and extension for 1 min/Kb. The mutagenesis of *RPGRIP1* exon 15 essential donor site was done using primers designed through agilent (http://www.genomics.agilent.com/primerDesignProgram.jsp). Site-directed mutagenesis was done using QuickChang II (Agilent Technologies, Santa Clara, CA) and the results confirmed via Sanger sequencing.

Quantitative polymerase chain reactions (qPCR). 5 ng of genomic DNA, 200 nM of each primer and 10 µl of Fast SYBR Green Master Mix (Life Technologies, Grand Island, NY) were used for qPCR reactions which was performed on Stratagene Mx3000P (Agilent Technologies, Santa Clara, CA) using the standard thermocycling program (95 °C for 3 min, 40 cycles of 95 °C for 20 s, and 60 °C for 1 min followed by a melting curve). Prior to the experiments, the primers used were evaluated for efficiency via standard curve generation of serially diluted control DNA and shown to have an efficiency of >90%. The ddCT method was used for the analysis of results where *ZNF80* was used as a reference gene and an in-house DNA sample with wild type *RPGRIP1* (OGI200) used for normalization. Each samples was done in triplicate

and the average value was used. Standard deviation with error propagation was used to calculate up and down errors which are presented as error bars.

Acknowledgments

This work was supported by grants from the National Eye Institute [RO1EY012910 (EAP), R01EY026904 (KMB/EAP) and P30EY014104 (MEEI core support)], and the Foundation Fighting Blindness (USA, EAP). Sequencing and analysis was provided by the Center for Mendelian Genomics at the Broad Institute of MIT and Harvard and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung and Blood Institute grant UM1 HG008900 to Daniel MacArthur and Heidi Rehm. The authors would like to thank the patients and their family members for their participation in this study and the Ocular Genomics Core members for their experimental assistance. The authors would like to thank the Exome Aggregation Consortium, the Genome Aggregation Database (gnomAD) and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about and http://gnomad.broadinstitute .org/about. The authors declare no conflict of interest related to the work presented in this manuscript.

References

1. Leber T. Ueber Retinitis pigmentosa und angeborene Amaurose. Graefe's Archive For Clinical And Experimental Ophthalmology. 1869;15(3):1-25.
2. Alström CH, Olson O. Heredo-retinopathia congenitalis: monohybrida recessiva autosomalis: Berlingska boktryckeriet; 1957.
3. Franceschetti A, Dieterle P. [Diagnostic and prognostic importance of the electroretinogram in tapetoretinal degeneration with reduction of the visual field and hemeralopia]. Confinia neurologica. 1954;14(2-3):184-6. PubMed PMID: 13190865. Importance diagnostique et pronostique de l'electroretinogramme (ERG) dans les degenerescences tapeto-retiniennes avec retrecissement du champ visuel et hemeralopie.
4. Schappert-Kimmijser J, Henkes HE, Van Den Bosch J. Amaurosis congenita (Leber). AMA archives of ophthalmology. 1959;61(2):211-8.
5. Kumaran N, Moore AT, Weleber RG, Michaelides M. Leber congenital amaurosis/early-onset severe retinal dystrophy: clinical features, molecular genetics and therapeutic interventions. British Journal of Ophthalmology. 2017;101(9):1147-54.
6. Chacon-Camacho OF, Zenteno JC. Review and update on the molecular basis of Leber congenital amaurosis. World Journal of Clinical Cases: WJCC. 2015;3(2):112.
7. Henderson RH, Williamson KA, Kennedy JS, Webster AR, Holder GE, Robson AG, et al. A rare de novo nonsense mutation in OTX2 causes early onset retinal dystrophy and pituitary dysfunction. Molecular vision. 2009;15:2442.
8. den Hollander AI, Jacoline B, de Kok YJ, van Soest S, van den Born LI, van Driel MA, et al. Mutations in a human homologue of Drosophila crumbs cause retinitis pigmentosa (RP12). Nature genetics. 1999;23(2):217-21.
9. Lee C, Wallingford JB, Gross JM. Cluap1 is Essential for Ciliogenesis and Photoreceptor Maintenance in the Vertebrate EyeCluap1 in the Vertebrate Eye. Investigative ophthalmology & visual science. 2014;55(7):4585-92.
10. Minegishi Y, Sheng X, Yoshitake K, Sergeev Y, Iejima D, Shibagaki Y, et al. CCT2 mutations evoke leber congenital amaurosis due to chaperone complex instability. Scientific reports. 2016;6:33742.
11. Sayer JA, Otto EA, O'Toole JF, Nurnberg G, Kennedy MA, Becker C, et al. The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4. Nature genetics. 2006;38(6):674-81.
12. Bellingham J, Wijesuriya SD, Evans K, Fryer A, Lennon G, Gregory CY. Genetic and physical localisation of the gene causing cone-rod dystrophy (CORD2). Degenerative Retinal Diseases: Springer; 1997. p. 295-302.
13. Sohocki MM, Malone KA, Sullivan LS, Daiger SP. Localization of retina/pineal-expressed sequences: identification of novel candidate genes for inherited retinal disorders. Genomics. 1999;58(1):29-33.
14. Bowne SJ, Sullivan LS, Blanton SH, Cepko CL, Blackshaw S, Birch DG, et al. Mutations in the inosine monophosphate dehydrogenase 1 gene (IMPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. Human molecular genetics. 2002;11(5):559-68.

15. Vincent A, Forster N, Maynes JT, Paton TA, Billingsley G, Roslin NM, et al. OTX2 mutations cause autosomal dominant pattern dystrophy of the retinal pigment epithelium. Journal of medical genetics. 2014;51(12):797-805.
16. Friedman JS, Chang B, Kannabiran C, Chakarova C, Singh HP, Jalali S, et al. Premature truncation of a novel protein, RD3, exhibiting subnuclear localization is associated with retinal degeneration. The American Journal of Human Genetics. 2006;79(6):1059-70.
17. Chiang P-W, Wang J, Chen Y, Fu Q, Zhong J, Chen Y, et al. Exome sequencing identifies NMNAT1 mutations as a cause of Leber congenital amaurosis. Nature genetics. 2012;44(9):972-4.
18. Dryja TP, Hahn LB, Kajiwara K, Berson EL. Dominant and digenic mutations in the peripherin/RDS and ROM1 genes in retinitis pigmentosa. Investigative ophthalmology & visual science. 1997;38(10):1972-82.
19. Farrar GJ, Jordan SA, Kenna P, Humphries MM, Kumar-Singh R, McWilliam P, et al. Autosomal dominant retinitis pigmentosa: localization of a disease gene (RP6) to the short arm of chromosome 6. Genomics. 1991;11(4):870-4.
20. Koenekoop RK, Wang H, Majewski J, Wang X, Lopez I, Ren H, et al. Mutations in NMNAT1 cause Leber congenital amaurosis and identify a new disease pathway for retinal degeneration. Nature genetics. 2012;44(9):1035-9.
21. Perrault I, Hanein S, Zanlonghi X, Serre V, Nicouleau M, Defoort-Delhemmes S, et al. Mutations in NMNAT1 cause Leber congenital amaurosis with early-onset severe macular and optic atrophy. Nature genetics. 2012;44(9):975-7.
22. den Hollander AI, Koenekoop RK, Mohamed MD, Arts HH, Boldt K, Towns KV, et al. Mutations in LCA5, encoding the ciliary protein lebercilin, cause Leber congenital amaurosis. Nature genetics. 2007;39(7):889-95.
23. Estrada-Cuzcano A, Koenekoop RK, Coppieters F, Kohl S, Lopez I, Collin RW, et al. IQCB1 mutations in patients with leber congenital amaurosis. Investigative ophthalmology & visual science. 2011;52(2):834-9.
24. Moore A. NMNAT1 mutations cause Leber congenital amaurosis. 2012.
25. Sergouniotis PI, Davidson AE, Mackay DS, Li Z, Yang X, Plagnol V, et al. Recessive mutations in KCNJ13, encoding an inwardly rectifying potassium channel subunit, cause leber congenital amaurosis. The American Journal of Human Genetics. 2011;89(1):183-90.
26. Xiao X, Mai G, Li S, Guo X, Zhang Q. Identification of CYP4V2 mutation in 21 families and overview of mutation spectrum in Bietti crystalline corneoretinal dystrophy. Biochemical and biophysical research communications. 2011;409(2):181-6.
27. Camuzat A, Dollfus H, Rozet J-M, Gerber S, Bonneau D, Bonnemaison M, et al. A gene for Leber's congenital amaurosis maps to chromosome 17p. Human molecular genetics. 1995;4(8):1447-52.
28. Payne AM, Morris AG, Downes SM, Johnson S, Bird AC, Moore AT, et al. Clustering and frequency of mutations in the retinal guanylate cyclase (GUCY2D) gene in patients with dominant cone-rod dystrophies. Journal of medical genetics. 2001;38(9):611-4.

29. Perrault I, Saunier S, Hanein S, Filhol E, Bizet AA, Collins F, et al. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. The American Journal of Human Genetics. 2012;90(5):864-70.
30. Abu-Safieh L, Alrashed M, Anazi S, Alkuraya H, Khan AO, Al-Owain M, et al. Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. Genome research. 2013;23(2):236-47.
31. Asai-Coakwell M, March L, Dai XH, DuVal M, Lopez I, French CR, et al. Contribution of growth differentiation factor 6-dependent cell survival to early-onset retinal dystrophies. Human molecular genetics. 2013;22(7):1432-42.
32. Dryja TP, Adams SM, Grimsby JL, McGee TL, Hong D-H, Li T, et al. Null RPGRIP1 alleles in patients with Leber congenital amaurosis. The American Journal of Human Genetics. 2001;68(5):1295-8.
33. Gerber S, Perrault I, Hanein S, Barbet F, Ducroq D, Ghazi I, et al. Complete exon-intron structure of the RPGR-interacting protein (RPGRIP1) gene allows the identification of mutations underlying Leber congenital amaurosis. European journal of human genetics: EJHG. 2001;9(8):561.
34. Athanassiadou A, Voutsinas G, Psiouri L, Leroy E, Polymeropoulos MH, Ilias A, et al. Genetic analysis of families with Parkinson disease that carry the Ala53Thr mutation in the gene encoding alpha-synuclein. American journal of human genetics. 1999;65(2):555.
35. Gu S, Lennon A, Li Y, Lorenz B, Fossarello M, North M, et al. Tubby-like protein-1 mutations in autosomal recessive retinitis pigmentosa. The lancet. 1998;351(9109):1103-4.
36. Janecke AR, Thompson DA, Utermann G, Becker C, Hübner CA, Schmid E, et al. Mutations in RDH12 encoding a photoreceptor cell retinol dehydrogenase cause childhood-onset severe retinal dystrophy. Nature genetics. 2004;36(8):850-4.
37. Morimura H, Fishman GA, Grover SA, Fulton AB, Berson EL, Dryja TP. Mutations in the RPE65 gene in patients with autosomal recessive retinitis pigmentosa or leber congenital amaurosis. Proceedings of the National Academy of Sciences. 1998;95(6):3088-93.
38. Wang H, den Hollander AI, Moayedi Y, Abulimiti A, Li Y, Collin RW, et al. Mutations in SPATA7 cause Leber congenital amaurosis and juvenile retinitis pigmentosa. The American Journal of Human Genetics. 2009;84(3):380-7.
39. Talib M, van Schooneveld MJ, van Genderen MM, Wijnholds J, Florijn RJ, Jacoline B, et al. Genotypic and phenotypic characteristics of CRB1-associated retinal dystrophies: a long-term follow-up study. Ophthalmology. 2017;124(6):884-95.
40. Li T. Leber congenital amaurosis caused by mutations in RPGRIP1. Cold Spring Harbor perspectives in medicine. 2015;5(4):a017384.
41. Boylan JP, Wright AF. Identification of a novel protein interacting with RPGR. Human molecular genetics. 2000;9(14):2085-93.
42. Hong D-H, Yue G, Adamian M, Li T. Retinitis pigmentosa GTPase regulator (RPGR)-interacting protein is stably associated with the photoreceptor ciliary axoneme and anchors RPGR to the connecting cilium. Journal of Biological Chemistry. 2001;276(15):12091-9.

43. Roepman R, Bernoud-Hubac N, Schick DE, Maugeri A, Berger W, Ropers H-H, et al. The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. Human molecular genetics. 2000;9(14):2095-105.
44. Mavlyutov TA, Zhao H, Ferreira PA. Species-specific subcellular localization of RPGR and RPGRIP isoforms: implications for the phenotypic variability of congenital retinopathies among species. Human molecular genetics. 2002;11(16):1899-907.
45. Castagnet P, Mavlyutov T, Cai Y, Zhong F, Ferreira P. RPGRIP1s with distinct neuronal localization and biochemical properties associate selectively with RanBP2 in amacrine neurons. Human molecular genetics. 2003;12(15):1847-63.
46. Shu X, Fry A, Tulloch B, Manson F, Crabb J, Khanna H, et al. RPGR ORF15 isoform co-localizes with RPGRIP1 at centrioles and basal bodies and interacts with nucleophosmin. Human Molecular Genetics. 2005;14(9):1183-97.
47. Zhao Y, Hong D-H, Pawlyk B, Yue G, Adamian M, Grynberg M, et al. The retinitis pigmentosa GTPase regulator (RPGR)-interacting protein: subserving RPGR function and participating in disk morphogenesis. Proceedings of the National Academy of Sciences. 2003;100(7):3965-70.
48. Pawlyk BS, Bulgakov OV, Liu X, Xu X, Adamian M, Sun X, et al. Replacement gene therapy with a human RPGRIP1 sequence slows photoreceptor degeneration in a murine model of Leber congenital amaurosis. Human gene therapy. 2010;21(8):993-1004.
49. Lhériteau E, Petit L, Weber M, Le Meur G, Deschamps J-Y, Libeau L, et al. Successful gene therapy in the RPGRIP1-deficient dog: a large model of cone–rod dystrophy. Molecular therapy. 2014;22(2):265-77.
50. Consugar MB, Navarro-Gomez D, Place EM, Bujakowska KM, Sousa ME, Fonseca-Kelly ZD, et al. Panel-based genetic diagnostic testing for inherited eye diseases is highly accurate and reproducible and more sensitive for variant detection than exome sequencing. Genetics in medicine: official journal of the American College of Medical Genetics. 2015;17(4):253.
51. Handsaker RE, Van Doren V, Berman JR, Genovese G, Kashin S, Boettger LM, et al. Large multiallelic copy number variations in humans. Nature genetics. 2015;47(3):296-303.
52. Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. Current opinion in cell biology. 2004;16(3):293-9.
53. Bujakowska KM, Fernandez-Godino R, Place E, Consugar M, Navarro-Gomez D, White J, et al. Copy-number variation is an important contributor to the genetic causality of inherited retinal degenerations. Genetics in medicine : official journal of the American College of Medical Genetics. 2017 Jun;19(6):643-51. PubMed PMID: 27735924.
54. Desmet F-O, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic acids research. 2009;37(9):e67-e.
55. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91.

56. Russell S, Bennett J, Wellman JA, Chung DC, Yu Z-F, Tillman A, et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. The Lancet. 2017;390(10097):849-60.
57. Zallocchi M, Binley K, Lad Y, Ellis S, Widdowson P, Iqball S, et al. EIAV-based retinal gene therapy in the shaker1 mouse model for usher syndrome type 1B: development of UshStat. PLoS One. 2014;9(4):e94272.
58. Banin E, Gootwine E, Obolensky A, Ezra-Elia R, Ejzenberg A, Zelinger L, et al. Gene augmentation therapy restores retinal function and visual behavior in a sheep model of CNGA3 achromatopsia. Molecular Therapy. 2015;23(9):1423-33.
59. Komáromy AM, Alexander JJ, Rowlan JS, Garcia MM, Chiodo VA, Kaya A, et al. Gene therapy rescues cone function in congenital achromatopsia. Human molecular genetics. 2010;19(13):2581-93.
60. Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, et al. Effect of gene therapy on visual function in Leber's congenital amaurosis. New England Journal of Medicine. 2008;358(21):2231-9.
61. MacLaren RE, Groppe M, Barnard AR, Cottriall CL, Tolmachova T, Seymour L, et al. Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial. The Lancet. 2014;383(9923):1129-37.
62. Maguire AM, Simonelli F, Pierce EA, Pugh Jr EN, Mingozzi F, Bennicelli J, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. New England Journal of Medicine. 2008;358(21):2240-8.
63. Koenekoop RK. An overview of Leber congenital amaurosis: a model to understand human retinal development. Survey of ophthalmology. 2004;49(4):379-98.
64. Miyagawa M, Naito T, Nishio SY, Kamatani N, Usami S. Targeted exon sequencing successfully discovers rare causative genes and clarifies the molecular epidemiology of Japanese deafness patients. PLoS One. 2013;8(8):e71381. PubMed PMID: 23967202. Pubmed Central PMCID: 3742761.
65. Audo I, Bujakowska KM, Leveillard T, Mohand-Said S, Lancelot ME, Germain A, et al. Development and application of a next-generation-sequencing (NGS) approach to detect known and novel gene defects underlying retinal diseases. Orphanet journal of rare diseases. 2012 Jan 25;7:8. PubMed PMID: 22277662. Pubmed Central PMCID: 3352121.
66. Glockle N, Kohl S, Mohr J, Scheurenbrand T, Sprecher A, Weisschuh N, et al. Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. European journal of human genetics : EJHG. 2014 Jan;22(1):99-104. PubMed PMID: 23591405. Pubmed Central PMCID: 3865404.
67. Coppieters F, De Wilde B, Lefever S, De Meester E, De Rocker N, Van Cauwenbergh C, et al. Massively parallel sequencing for early molecular diagnosis in Leber congenital amaurosis. Genetics in medicine : official journal of the American College of Medical Genetics. 2012 Jun;14(6):576-85. PubMed PMID: 22261762.
68. Neveling K, Collin RW, Gilissen C, van Huet RA, Visser L, Kwint MP, et al. Next-generation genetic testing for retinitis pigmentosa. Human mutation. 2012 Jun;33(6):963-72. PubMed PMID: 22334370. Pubmed Central PMCID: 3490376.

69. Eisenberger T, Neuhaus C, Khan AO, Decker C, Preising MN, Friedburg C, et al. Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: the example of retinal dystrophies. PLoS One. 2013;8(11):e78496. PubMed PMID: 24265693. Pubmed Central PMCID: 3827063.
70. Van Cauwenbergh C, Van Schil K, Cannoodt R, Bauwens M, Van Laethem T, De Jaegere S, et al. arrEYE: a customized platform for high-resolution copy number analysis of coding and noncoding regions of known and candidate retinal dystrophy genes and retinal noncoding RNAs. Genetics in medicine : official journal of the American College of Medical Genetics. 2017 Apr;19(4):457-66. PubMed PMID: 27608171. Pubmed Central PMCID: 5392597.
71. Ferrari S, Di Iorio E, Barbaro V, Ponzin D, S Sorrentino F, Parmeggiani F. Retinitis pigmentosa: genes and disease mechanisms. Current genomics. 2011;12(4):238-49.
72. Remans K, Bürger M, Vetter IR, Wittinghofer A. C2 domains as protein-protein interaction modules in the ciliary transition zone. Cell reports. 2014;8(1):1-9.
73. Carss KJ, Arno G, Erwood M, Stephens J, Sanchis-Juan A, Hull S, et al. Comprehensive Rare Variant Analysis via Whole-Genome Sequencing to Determine the Molecular Pathology of Inherited Retinal Disease. Am J Hum Genet. 2017 Jan 05;100(1):75-90. PubMed PMID: 28041643. Pubmed Central PMCID: 5223092.
74. Yeo G, Holste D, Kreiman G, Burge CB. Variation in alternative splicing across human tissues. Genome biology. 2004;5(10):R74.
75. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Science translational medicine. 2017;9(386):eaal5209.
76. Kimberling WJ. Estimation of the frequency of occult mutations for an autosomal recessive disease in the presence of genetic heterogeneity: application to genetic hearing loss disorders. Human mutation. 2005;26(5):462-70.
77. Falk MJ, Zhang Q, Nakamaru-Ogiso E, Kannabiran C, Fonseca-Kelly Z, Chakarova C, et al. NMNAT1 mutations cause Leber congenital amaurosis. Nature genetics. 2012;44(9):1040-5.
78. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nature biotechnology. 2011;29(1):24-6.
79. Farkas MH, Grant GR, White JA, Sousa ME, Consugar MB, Pierce EA. Transcriptome analyses of the human retina identify unprecedented transcript diversity and 3.5 Mb of novel transcribed sequence via significant alternative splicing and novel genes. BMC genomics. 2013;14(1):486.

ID	Age	Sex	Onset	Variant 1	Variant 2	Symptoms	VA	Signs	NGS
OGI281-608	61	M	Early Childhood	c.1615_1624del10 (p.E539Qfs*2) [0.000004]	Exon2Dup	↓ Vision, Nyctalopia, Photophobia, ↓ CV	20/400	↓VF, ↓ ERG, Nystagmus, ON atrophy, Attenuated BV, Bone Spicules	GEDi, WES WGS
OGI237-523	50	F	Birth	c.3793ins4 (p.V1265Gfs*19) [0.000004]	Exon1-2Dup	↓ Vision, Light Sensitivity, ↓ CV	20/500	↓VF, ↓ ERG, Keratoconus, PSC, Asteroid Hyalosis, Attenuated BV, Bone Spicules	GEDi WES
OGI827-1591	30	M	1.5 yo	c.895_896del2 (p.G299Sfs*21) [0.0]	c.2367+23delG [0.00023]	↓ Vision	LP	↓VF, ↓ ERG, Nystagmus, Macular atrophy, Attenuated BV, Bone Spicules	GEDi WES
OGI1797_3128	N/A	N/A	N/A	c.2302C>T (p.R768*) [0.00002]	Exon19del	N/A	N/A	N/A	GEDi
OGI501-336	N/A	N/A	N/A	c.2302C>T (p.R768*) [0.00002]	c.711_711del1 (p. P237Pfs*40) [0.0]	N/A	N/A	N/A	GEDi
OGI690-1378	N/A	N/A	N/A	c.1084_1087del (p.E362NAfs*12) [0.0]	c.767C>G p.S256* [0.0]	N/A	N/A	N/A	GEDi
OGI0079_194	N/A	N/A	N/A	c.3793ins4 (p.V1265Gfs*19) [0.000004]	c.3793ins4 (p.V1265Gfs*19) [0.000004]	N/A	N/A	N/A	GEDi

Table 1. Clinical information and mutations list. The clinical features and *RPGRIP1* mutations of the probands of the families studied. There was no consanguinity in any of the families. The transcript used is RefSeq NM_020366. The frequency of the alleles reported in gnomAD data base (55), are written in brackets under the variants. Dup= Exon duplication, Del= exon deletion, CV= color vision, VA= best visual acuity, LP= light perception, VF= visual filed, ERG= electroretinogram signal, ON= optic nerve, BV= blood vessels, NGS= next-generation sequencing method, GEDi= genetic eye diseases targeted sequencing(50), WES= whole exome sequencing, WGS= whole genome sequencing.

Figure 1. Clinical findings. Typical clinical findings of the patients described in this paper, including fundus images of the right (A) and the left (B) eyes of OGI827-1591 (c.895-896del2(p.G299Sfs*21), c.2367+23delG) showing bone spicules. Right eye fundus image of OGI281-608 (c.1615-1624del10 (p.E539QfsTer2), allele 2: Exon2Dup) is shown in (C). OCT of the same eye shows sparing of the macula (D). This patient had extensive disease with similar findings in the left eye.

Figure 2. Outline of the sequencing approach. A total of 1719 families underwent GEDi sequencing. A subgroup of 3 families had the two defective alleles of *RPGRIP1* identified and 4 had one allele detected. These families had the second allele identified through a step–wise expansion of analysis from copy number studies, to exome and genome sequencing. The number of families in each category are shown in bold text.

Figure 3. The structural mutation in OGI281. The exon 2 tandem duplication was

shown by WGS reads on the integrative genome viewer (igv) (78) (A) The change in orientation of the paired end reads enabled prediction of the breakpoint (B). The duplication was subsequently validated by polymerase chain reaction (PCR) across the breakpoint (C) and subsequent Sanger sequencing (black arrow head) (D). The band used for Sanger sequencing is highlighted with a red star in C.

Figure 4. Copy number validation. Three families had copy number mutations in *RPGRIP1*. The quantitative polymerase chain reaction (qPCR) results are shown in this figure. OGI281 and OGI237 had a gain of copy in exon 2 and exons 1 and 2 respectively while OGI1797_003128 had a deletion of exon 19. The genomic coordinates are shown on the X-axis and the corresponding exon areas are heighted in blue. All coordinates in this paper are based on hg38.

Figure 5. Analysis of the deep intronic mutation in family OGI827. The pedigree of the family is shown in A revealing the segregation of the two *RPGRIP1* mutations. The c.2367+23delG and wild type (WT) alleles were cloned from the proband's father. Site directed mutagenesis on the WT allele was used to create a c.2367+1G>A as positive control for a splicing defect in intron 15. After transfection into HEK293t cells and complementary DNA generation, reverse transcription polymerase chain reaction (RT-PCR) reveals a prominence of the unspliced transcript with c.2367+23delG (B). Sanger sequencing of the band highlighted by the red star in B, shows the expected retention of intron 15 while intron 14 is spliced out. The red arrowhead corresponds to the essential splice site mutation c.2367+1G>A.

Figure 6. *RPGRIP1* mutations. Novel coding mutations in *RPGRIP1* discovered

through our study are labeled in red along the *RPGRIP1* gene model and previously reported mutations in the Human Gene Mutation Database (HGMD) are in black. The structural/non-coding mutations are presented below the *RPGRIP1* model. Exons are represented as rectangles and introns as arrowed lines connecting the exons. The coordinates are described relevant to NM_020366 which is the transcript RefSeq transcript illustrated. The genomic distances are to scale horizontally and the genomic coordinated are based on hg38. Below the *RPGRIP1* gene, the protein with its various domains is shown and the associated exons are connected via gray highlight. ND: nuclear domain, CC: coiled coil domains, C2: protein kinase C conserved domain 2, RID: RPGR interacting domain.

Figure 1.

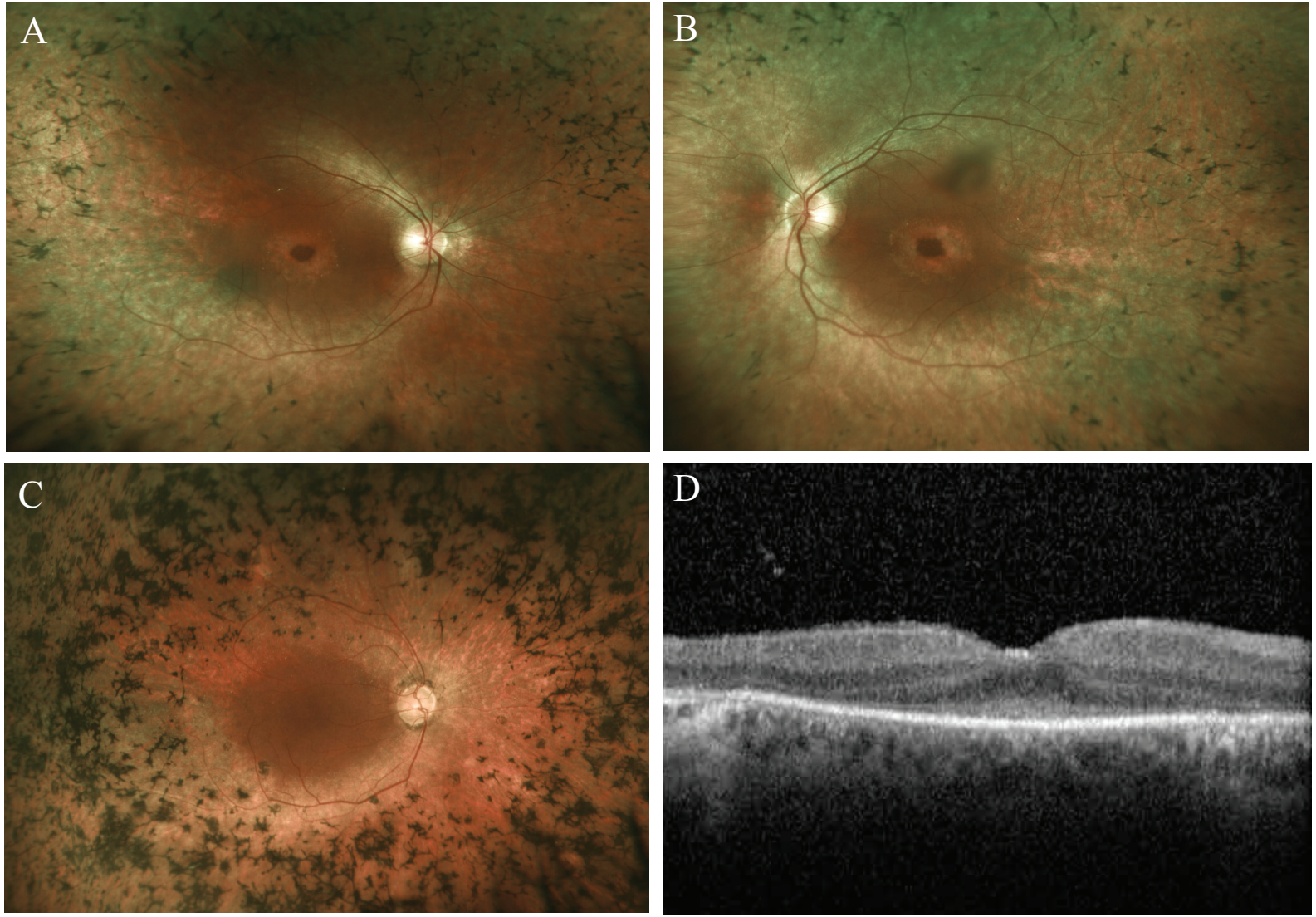


Figure 2.

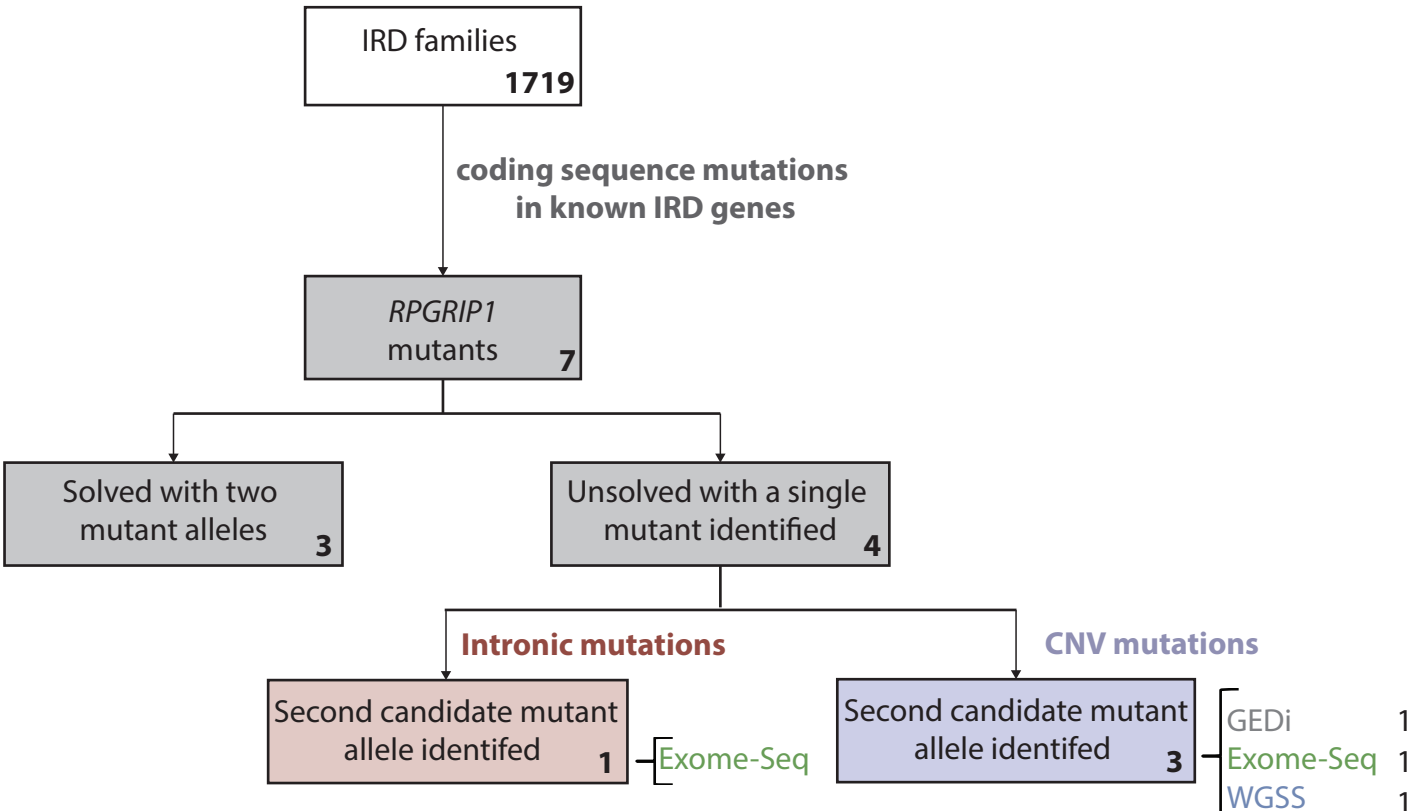
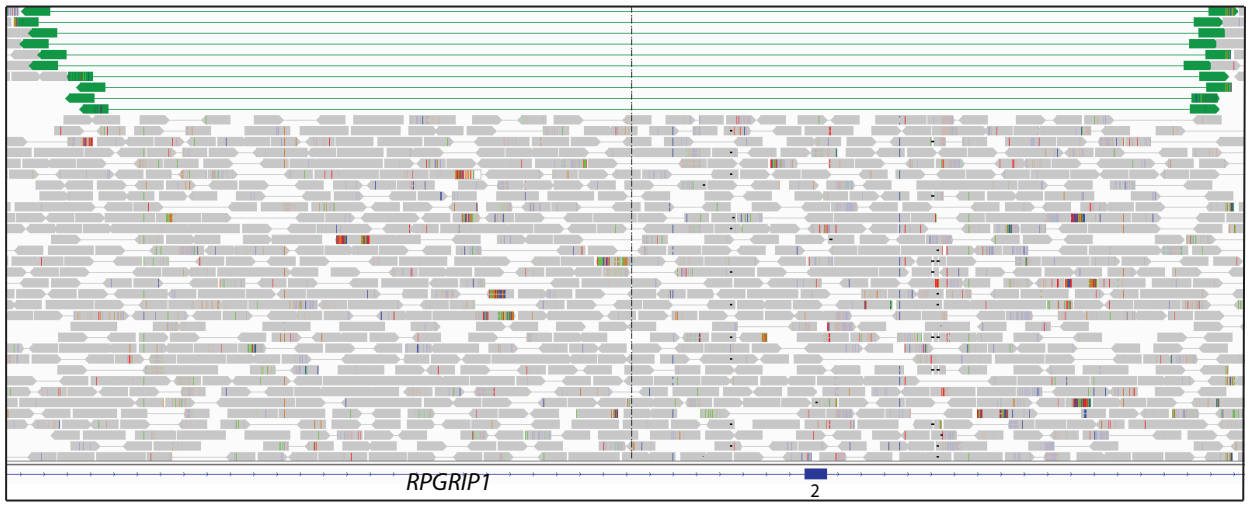
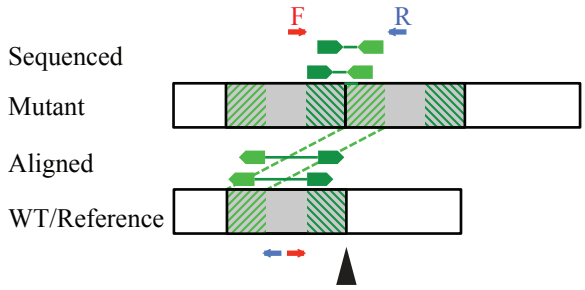

IRD families
1719
coding sequence mutations
in known IRD genes
RPGRIP1
mutants
7
Solved with two
mutant alleles
3
Unsolved with a single
mutant identified
4
Intronic mutations
CNV mutations
Second candidate mutant
allele identifed
1
Exome-Seq
Second candidate mutant
allele identifed
3
GEDi
1
Exome-Seq
1
WGSS
1

Figure 3.

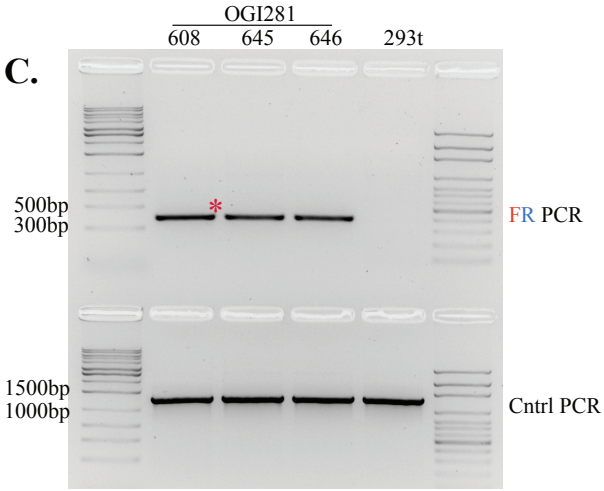
A.



B.



C.



D.

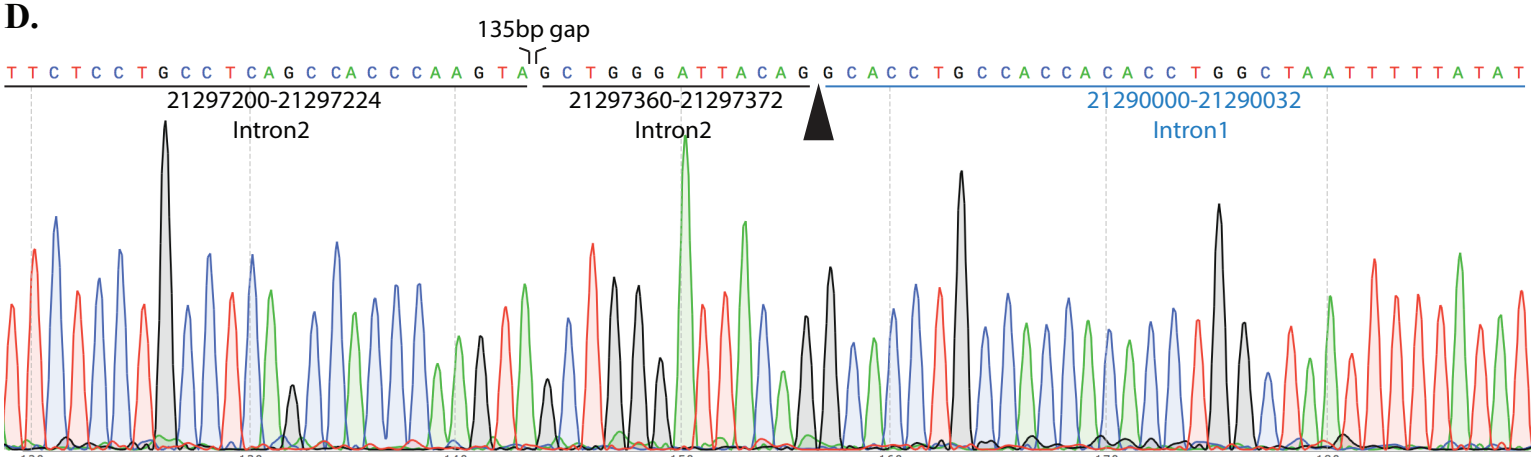


Figure 4.

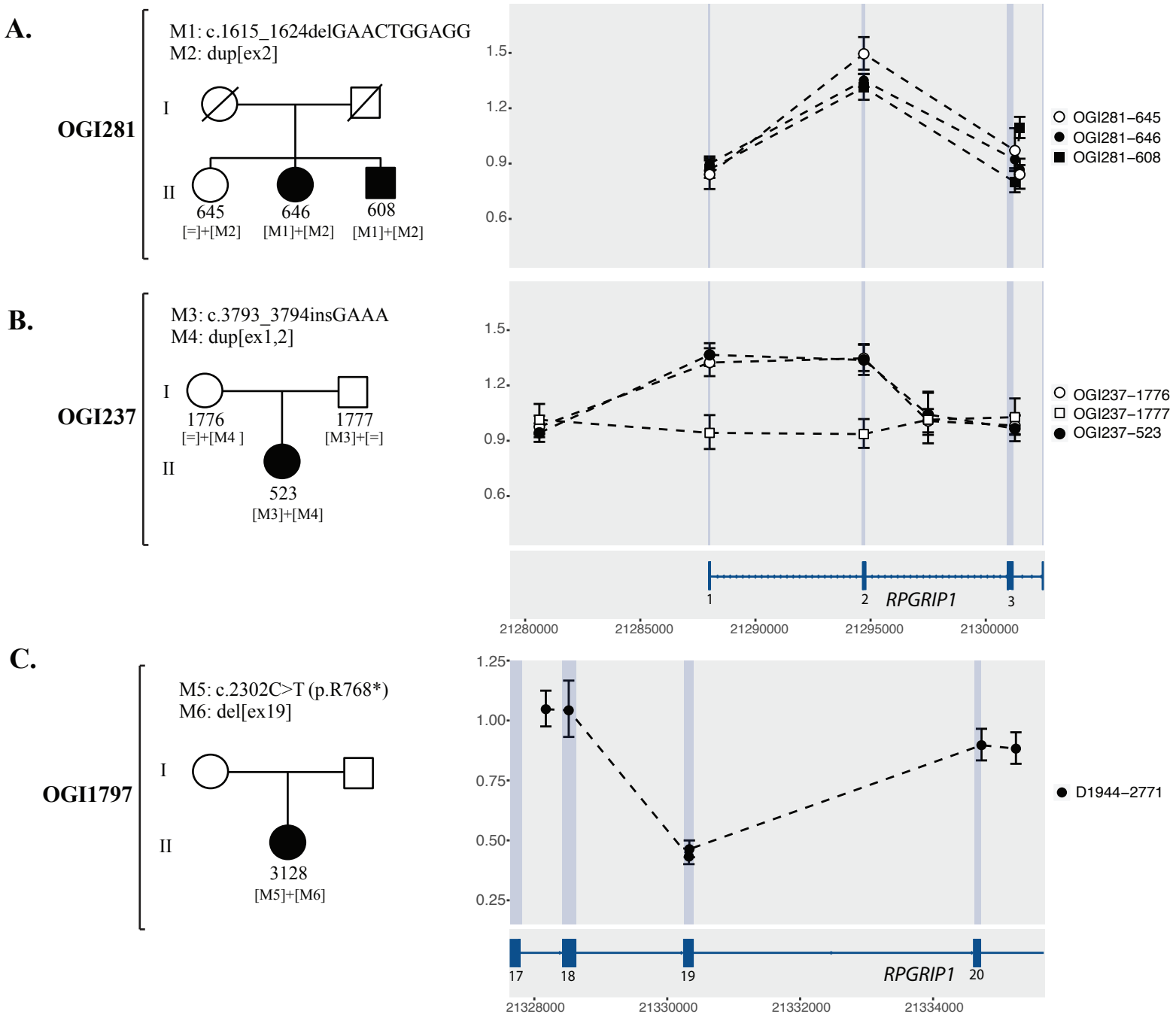
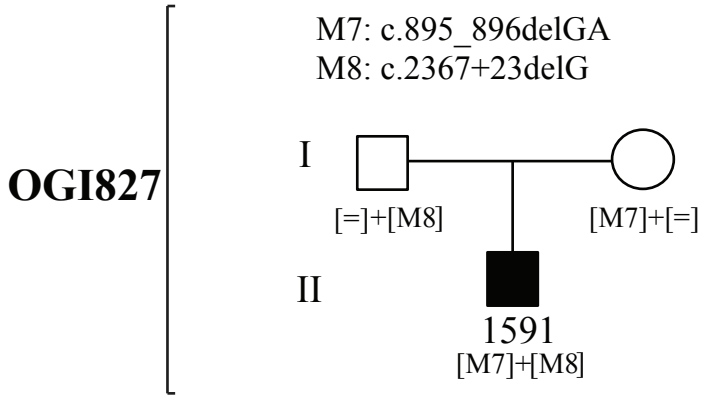
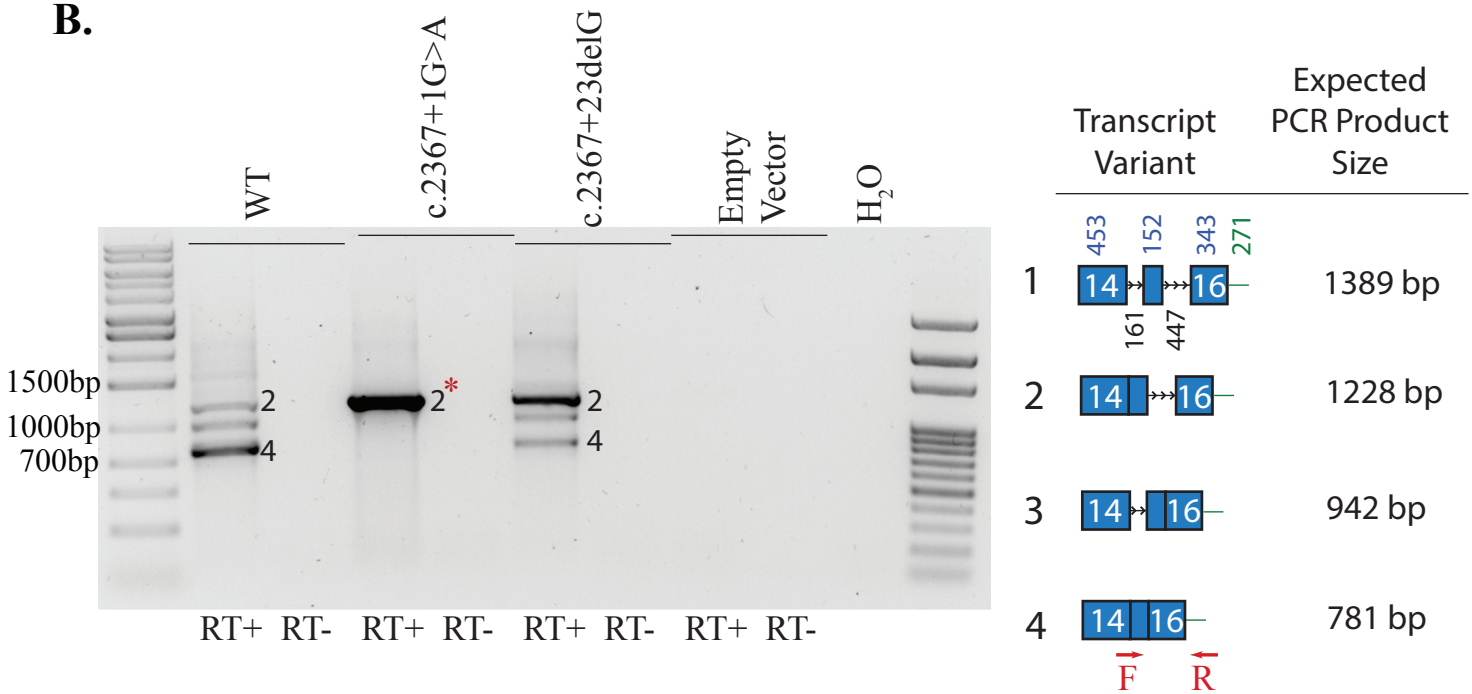

A.
OGI281
M1: c.1615_1624delGAACTGGAGG
M2: dup[ex2]
I
II
645
[=]+[M2]
646
[M1]+[M2]
608
[M1]+[M2]
1.5
1.2
0.9
0.6
OGI281–645
OGI281–646
OGI281–608
B.
OGI237
M3: c.3793_3794insGAAA
M4: dup[ex1,2]
I
II
1776
[=]+[M4]
1777
[M3]+[=]
523
[M3]+[M4]
1.5
1.2
0.9
0.6
OGI237–1776
OGI237–1777
OGI237–523
1
2
RPGRIP1
3
21280000
21285000
21290000
21295000
21300000
C.
OGI1797
M5: c.2302C>T (p.R768*)
M6: del[ex19]
I
II
3128
[M5]+[M6]
1.25
1.00
0.75
0.50
0.25
D1944–2771
17
18
19
RPGRIP1
20
21328000
21330000
21332000
21334000

Figure 5.

A.


OGI827
M7: c.895_896delGA
M8: c.2367+23delG
I
[=]+[M8]
[M7]+[=]
II
1591
[M7]+[M8]

B.


WT
c.2367+1G>A
c.2367+23delG
Empty Vector
H_2O
1500bp
1000bp
700bp
2
4
2*
2
4
RT+ RT- RT+ RT- RT+ RT- RT+ RT-
Transcript Variant
Expected PCR Product Size
453
152
343
271
161
447
1
14
16
1389 bp
2
14
16
1228 bp
3
14
16
942 bp
4
14
16
781 bp
F
R

C.

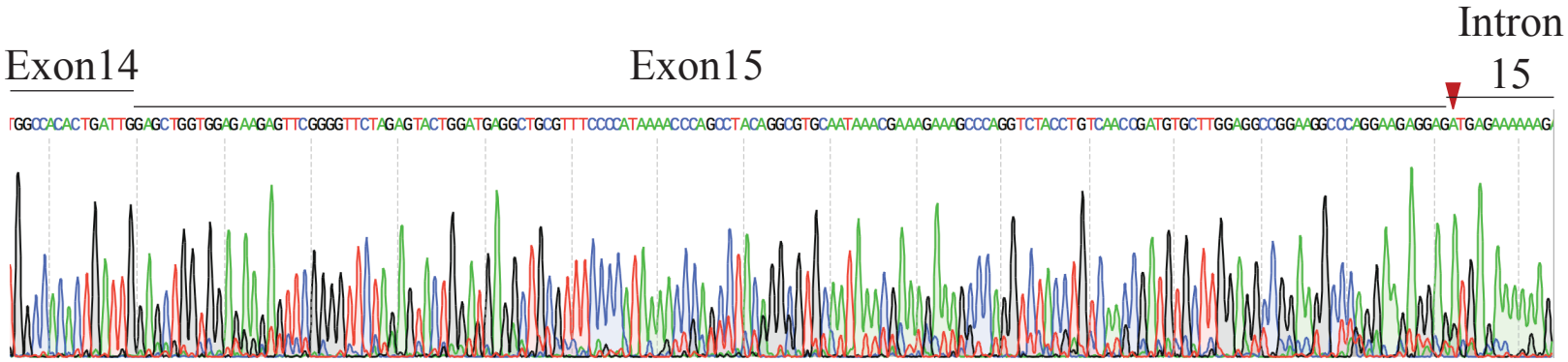
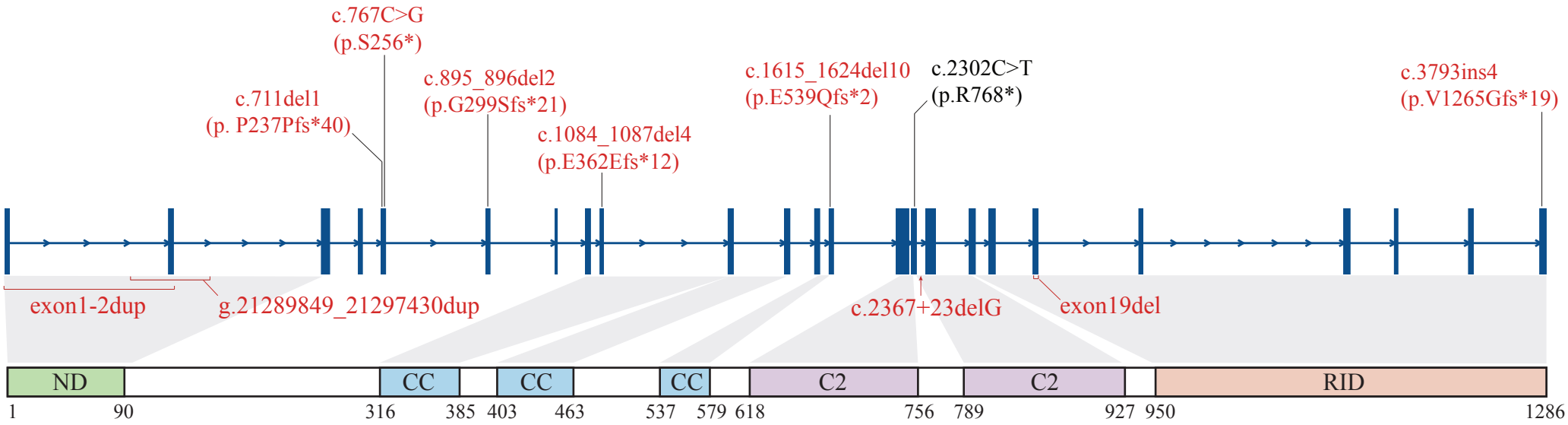

Exon14
Exon15
Intron 15

Figure 6.


c.767C>G
(p.S256*)
c.711del1
(p. P237Pfs*40)
c.895_896del2
(p.G299Sfs*21)
c.1084_1087del4
(p.E362Efs*12)
c.1615_1624del10
(p.E539Qfs*2)
c.2302C>T
(p.R768*)
c.3793ins4
(p.V1265Gfs*19)
exon1-2dup
g.21289849_21297430dup
c.2367+23delG
exon19del
ND
CC
CC
CC
C2
C2
RID
1
90
316
385
403
463
537
579
618
756
789
927
950
1286