Novel missense mutation E585K in retinitis pigmentosa leads to compromised *RPGR* splicing diversity

Yan-Shan Liu^{1,2,3#}, Ji-Feng Wan^{4#}, Chun-Yan Ren^{1,5#}, Zhou-Heng Xu¹, Xu-Bin Pan⁶, Jia-Qi Pan¹, Ruo-Nan Gao¹, Shao-Qiang Liu¹, Jia-Li Zhang¹, Qian-Hao Yao⁷, Ji-Hong Wang⁶, En-Min Li⁸, Jun-Hua Rao^{2,3}, Ping Hou^{7*}, Jian-Huan Chen^{1,2,3*}

¹Laboratory of Genomic and Precision Medicine, Wuxi School of Medicine, Jiangnan University, Wuxi, Jiangsu, China.

²Joint Primate Research Center for Chronic Diseases, Jiangnan University and Guangdong Institute of Applied Biological Resources, Jiangnan University, Wuxi, Jiangsu, China.

³Guangdong Institute of Applied Biological Resources, Guangzhou, Guangdong, China.

⁴Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangzhou, China.

⁵ School of Biotechnology, Wuxi, Jiangsu, China.

⁶Affiliated Hospital of Jiangnan University, Wuxi, Jiangsu, China.

⁷Jinhua Eye Hospital, Jinhua, Zhejiang, China

⁸Shantou University Medical College, Shantou, Guangdong, China

Running title: ***RPGR* splicing diversity in retinitis pigmentosa**

[#]These authors contributed equally to this study.

^{*}Corresponding authors:

Ping Hou, Jinhua Eye Hospital, E-mail address: hou_33@163.com.

Jian-Huan Chen, Laboratory of Genomic and Precision Medicine, Wuxi School of Medicine,

Jiangnan University. E-mail address: cjh_bio@hotmail.com.

Abstract

Mutations in the retinitis pigmentosa GTPase regulator (*RPGR*) gene, are the major cause of X-linked retinitis pigmentosa (RP). Herein we used whole-exome sequencing to screen possible novel *RPGR* mutations in RP patients, and identified a novel missense mutation E585K in a patient with early onset but slow disease progression, and a frameshift deletion E998Gfs*78 in a patient with RP sine pigmento and high myopia. Intriguingly, bioinformatic analysis indicated that E585K probably affected *RPGR* RNA splicing instead of the protein sequence directly. Mini-gene assays in 293T cells revealed that splicing events of the E585K mutant were found to be also exist in wildtype, but with a shifted pattern. In the E585K mini-gene usage of an upstream alternative 5' splice site (5' ss) of exon 14 was enhanced, and other splicing events were suppressed, including the canonical 5' ss of exon 14, skipping of exon 14/15 and retention of intron 14. As a result, *RPGR* splicing products of the E585K mini-gene were predominated by transcripts containing a 4-bp deletion, with a small fraction of in-frame transcripts containing a retained intron 14, which might explain the slow disease progression in the patient carrying the mutation. RNA-Seq analysis further confirmed existence of these splicing events in endogenous *RPGR* RNA in human retina, pointing to compromised splicing diversity in the E585K mutant. Our findings thus added to the understanding of genotype-phenotype correlation in RP, and suggested that compromised *RPGR* splicing diversity might play a role in molecular mechanism of the disease.

Keywords: Retinitis pigmentosa, *RPGR*, novel mutations, alternative splicing

Introduction

Retinitis pigmentosa (RP, MIM 268000) is a group of inherited retinal degenerative diseases characterized by night blindness due to the progressive loss of rod photoreceptors, followed by complete blindness due to the further loss of cones (Hamel 2006, Tsang and Sharma 2018). X linked retinitis pigmentosa (XLRP) represents the most severe and prevalent form owing to its early onset and severity of progress (Tee et al. 2018, Tsang and Sharma 2018). To date, six loci have been associated with XLRP, yet only three genes have been identified (Lyraki et al. 2016, Ferrante et al. 2001, Schwahn et al. 1998, Meindl et al. 1996), among which the retinitis pigmentosa GTPase regulator (*RPGR*) gene [MIM 312610] accounts for 70-90% of XLRP and 15-20% of simplex RP cases, making it one of the most common causes of RP (Parmeggiani et al. 2017, Khanna 2018).

RPGR locates at Xp11.4 and undergoes complex alternative splicing in its RNA expression. Two major transcripts, the constitutive *RPGR*_{ex1-19} and *RPGR*_{ORF15}, have been extensively studied in XLRP (Megaw et al. 2015a, Charng et al. 2019). *RPGR*_{ORF15} shares the same exon 1-14 with *RPGR*_{ex1-19} and uses an alternative large ORF15 exon. *RPGR*_{ORF15} encodes a purine-rich, highly repetitive sequence followed by a basic C terminal domain with unknown function (Patnaik et al. 2015). The *RPGR* ORF15 region is also a mutation hotspot, accounting for around 60% of *RPGR* pathogenic mutations (Beltran et al. 2014, Vervoort et al. 2000). Most of the mutations in this region are one to two, or four to five bp deletions in exon, resulting in a truncated protein with the loss of C terminal domain (Cehajic Kapetanovic et al. 2019). The Constitutive *RPGR*_{ex1-19} is widely expressed while *RPGR*_{ORF15} is highly expressed in retina. Overexpression of *RPGR*_{ex1-19} results in severe retinal degeneration, suggesting a delicate ratio

between two isoforms, regulated by alternative splicing, is important for integrity and function of retina (Liu and Zack 2013, Wright et al. 2011).

In the current study, our whole-exome sequencing (WES) in male RP patients identified two mutations in *RPGR* (NM_001034853), including a novel missense mutation c.1753G>A (E585K) in a patient with early onset but slow disease progression, and a frameshift deletion E998Gfs*78 (Vervoort et al. 2000) in a patient with RP sine pigmento and high myopia. The E585K mutant mini-gene showed compromised *RPGR* splicing diversity, resulting in predominant frameshift transcripts and a small fraction of transcripts with intron 14 retended. It was noteworthy that all of these splicing variants were also found in endogenous *RPGR* expression across multiple human cell types, but the compositions of splicing variants were altered compared to wildtype. Our findings indicated that compromised splicing diversity instead of biogenesis of new disease-causing transcripts in *RPGR* expression could play a potential role in molecular mechanism of RP.

Methods

Clinical examination and sample collection

This study was approved by the Ethics Committee of Jiangnan University, and was conducted in accordance with the Declaration of Helsinki. Written consent was obtained from all participants after explanation of the nature of the study. Eleven male unrelated Chinese nonsyndromic RP patients were included in the current study after comprehensive ophthalmic examination. Peripheral blood was collected from all of the patients, and genomic DNA was extracted using TIANamp Blood DNA Maxi Kit (TIANGEN, Beijing, China).

87 *Whole exome sequencing*

88 Genomic DNA (3 µg per individual) from selected individuals was taken for WES by
89 Novogene (Tianjin, China). The whole exome was captured by an Agilent SureSelect All
90 Human exon v5.0 kit (Agilent, Santa Clara, US.), and sequenced on an Illumina Hiseq X-10
91 (Illumina, Hayward, US.) with a paired-end 150 bp length configuration.

92 *Analysis of whole exome sequencing data*

93 The WES data were analyzed as previously described (Chen et al. 2016). Quality of raw
94 sequencing data was analyzed by using FastQC
95 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). BWA (Burrows-Wheeler
96 Aligner, Version 0.7.6) was used for mapping short reads against the Human Reference
97 Genome hg19 from UCSC Genome Browser (<http://genome.ucsc.edu/>)(Li and Durbin 2009).
98 Variants were called and recalibrated according to the best practice pipeline by GATK (version
99 3.6), and then annotated using ANNOVAR (Wang et al. 2010). Minor allele frequencies (MAFs)
100 of genetic variants were estimated using available data from dbSNP138
101 (<http://www.ncbi.nlm.nih.gov/snp/>), the 1000 Genomes Project
102 (<http://www.1000genomes.org>), the NHLBI Exome Sequencing Project (ESP,
103 <http://evs.gs.washington.edu/EVS/>), the exome array design data
104 (http://genome.sph.umich.edu/wiki/Exome_Chip_Design) (Huyghe et al. 2013), and the
105 Exome Aggregation Consortium (ExAC, <http://exac.broadinstitute.org/>) and gnomAD
106 (<https://gnomad.broadinstitute.org/>) data. SIFT (<http://genetics.bwh.harvard.edu/pph/>) (Kumar
107 et al. 2009) and PolyPhen (http://sift.jcvi.org/www/SIFT_chr_coords_indels_submit.html)

(Ramensky et al. 2002) programs were used to predict the functional impact of remaining variants on the encoded protein. The WES detected variants were then screened for possible mutations in known genes resulting in retinitis pigmentosa according to Retinal Information Network (RetNet, <https://sph.uth.edu/retnet/>).

Sanger sequencing

The sequences of genes with selected variants were obtained from the NCBI reference sequence database (<http://www.ncbi.nlm.nih.gov/Refseq>). Primers were then designed to confirm selected variants; detailed information of primers was shown on Table 1. Primers were designed accordingly by Primer 3 are summarized in **Table S1**. Genomic DNA from patients was amplified by Polymerase chain reaction (PCR) using PrimeSTAR HS Premix (Takara Bio) on a Mastercycler nexus (Eppendorf, Hamburg, Germany) as described previously (Chen et al. 2016). Sequence alignment and analysis of variations were performed using the NovoSNP program (Weckx et al. 2005).

In-house control exomes

A dataset of 150 in-house control exomes, which was consolidated from multiple previous WES studies performed by our laboratory, was used to filter out variants that were not related to the disease phenotype. all control individuals had no sign or family history of retinitis pigmentosa and other inherited retinal diseases.

Mini-gene plasmid construction

To find out whether the presumed missense mutation affected *RPGR* splicing, a *RPGR* mini-gene was constructed based on pEGFP-N3 (TAKARA Bio, Kusatsu, Japan). Genomic

sequence spanning from exon13 to exon 15, including intron 13 and 14 of *RPGR* containing the wildtype and mutant of c.1753G>A were amplified from genomic DNA of an unrelated healthy control and patient 1 respectively, and inserted together with a Kozak sequence into the N terminus of EGFP. The *RPGR* mini-genes were expected to generate a fusion protein with EGFP.

Cell culture and transfection

Human 293T cells and ARPE-19 cells (American Tissue Culture Collection, Manassas, US.) were cultured in 6-well plates (Corning, Corning, US.) in Dulbecco's modified Eagle's medium: Ham's F-12 (DMEM/F12, ThermoFisher Scientific, Waltham, US.) supplemented with 10% fetal bovine serum (FBS, ThermoFisher Scientific, Waltham, US.) and antibiotics (100 units/ml penicillin G and 100 µg/ml streptomycin sulfate, ThermoFisher Scientific, Waltham, US.) at 37°C in a 5% CO₂ incubator. 293T cells were transfected with 2.5 µg of wild-type or mutant mini-gene plasmids using Lipofectamine 3000 (ThermoFisher Scientific, Waltham, USA) following the manufacturer's instruction.

Reverse transcription and PCR

Forty-eight hours after transfection total RNA was extracted using RNeasy Tissue & Cell kit (QIAGEN, Beijing, China), and reverse transcription (RT) was performed using PrimeScript RT reagent Kit with gDNA Eraser (TAKARA Bio, Kusatsu, Japan). Primers were designed to specifically amplify cDNA from transfected vectors and endogenous *RPGR* from 293T and ARPE-19 cells respectively. The sequences of PCR products were confirmed using Sanger sequencing. Primers for expression analysis are listed in **Table S1**.

Immunoblotting

Cells were lysed and total protein was extracted using protein lysis buffer (Biyotime, Shanghai, China) supplemented with a protease inhibitor (CWBIO, Beijing, China). Protein samples were run on a 12% acrylamide gel, and immunoblotting was performed using antibodies described above. Antibodies used in this study include mouse-anti-EGFP (AE012, ABClonal, Wuhan, China); HRP-conjugated mouse monoclonal anti-GAPDH (KC-5G5, Kangchen, Shanghai, China).

Splicing prediction

Splicing prediction of wild type and mutant variants was conducted using Human Splicing Finder (HSF, <http://www.umd.be/HSF/>) (Desmet et al. 2009).

Analysis of published retina RNA-Seq data

We used published RNA-seq data to investigate of *in vivo* *RPGR* splicing (Ratnapriya et al. 2019). Original data was downloaded from Gene Expression Omnibus (SRA, www.ncbi.nlm.nih.gov/geo) under access No. GSE115828. Original RNA-seq reads were aligned to human whole genome (hg38) with STAR (v2.7.0) (Dobin and Gingeras 2015), and visualized with IGV (v 2.4.17) (Robinson et al. 2011).

Results

WES identified two RPGR mutations

After analysis of exome sequencing data from the eleven male RP patients, we found two *RPGR* mutations in two RP patients. The summary of original exome sequencing data and detected variants of the two RP patients with *RPGR* mutations were shown in **Table S2** and

S3. The two hemizygotic *RPGR* mutations in two unrelated male Chinese RP patients, after exclusion of any possible mutation in other known RP causative genes (**Figure 1A**). Patient 1 had a novel missense mutation, c.1753G>A (E585K) and Patient 2 had a reported five-bp frameshift deletion, c.2993_2997del (E998Gfs*78) (Vervoort et al. 2000). The novel missense mutation was not found in either the public databases listed in the methods, or our control exomes. Sanger sequencing further confirmed both mutations in these two patients (**Figure 1B and 1C**). Multiple sequence alignment showed that for E585K, the mutated glutamine residue at codon 585 was not evolutionarily conserved (**Figure 1D**), and SIFT and Polyphen predicted the mutation to be benign (**Table 1**). In addition, the mutated nucleotide at 1753 position of the cDNA was not conserved among mammals, either (**Figure 1E**). However, c.1753G>A was at the last base of exon 14, one bp upstream to the canonical splice donor (5'ss), and thus might had potential impact on RNA splicing.

Clinical phenotyping results of the RP patients with RPGR mutations

Clinical phenotyping analysis was then conducted in the two RP patients with *RPGR* mutations (**Table 2**). Patient 1 with the E585K missense mutation was a 50-year-old male, with self-reported bilateral night blindness and blurry vision that started at around 10 years old. His symptoms progressed slowly after disease-onset, but deteriorated significantly over the past ten years. He had no other genetic eye disease or syndrome except posterior subcapsular cataracts, with normal intraocular pressure (< 21 mmHg). Fundus examination revealed bilateral severe, diffuse, granular appearance and atrophy of retinal pigment epithelium (RPE), and bone-spicule pigmentation, as well as arteriolar attenuation in mid-peripheral retina. Furthermore, optical coherence tomography (OCT) showed mild thinning of inner retinal layer

and severe thinning of outer retinal layer (**Figure 2A and 2B**).

Patient 2 was a 20-year-old male proband (III-3) from a three-generation XLRP family (**Figure S1A**), in which his younger brother (III-4) and one of his cousins (I-1) were also affected by RP. He suffered from gradually decreased visual acuity starting at six years old, with night blindness started from seven to eight years old. Fundus examination showed that he had bilateral unrecordable electroretinogram (ERG), but lacked characteristic bone spicule pigmentation, and thus was diagnosed as RP sine pigmento (**Figure S1B and S1C**). No other ocular disease or syndrome was found except high myopia.

Splicing analysis revealed splicing diversity near RPGR exon 14

To better understand the splicing of near the *RPGR* Exon 14, we then checked *RPGR* mRNA expression in human 293T cells and ARPE-19 cells (**Figure 3A**). Our Sanger sequencing results indicated that in both cell lines, there were probably two distinct 5'ss in (**Figure 3B**). In addition to a canonical 5'ss at the boundary of exon 14 (E14c), a novel alternative 5'ss was found at 4 bp upstream of it. As a result, splicing led to a 4 base frameshift deletion and a shorter exon 14 (E14s). Intriguingly, it was noted that the relative level of the alternative 5'ss usage and E14s was lower in ARPE-19 cells compared to 293T cells, probably in line with a higher expression level of *RPGR*_{ORF15} in the retina. Furthermore, by using published human retina RNA-seq data, we confirmed that retina expressed transcripts that lacked the last 4 bp of exon 14 (**Figure S2**), which further supported the existence of the intrinsic alternative 5'ss and E14s in normal human retina. In addition, RT-PCR followed by Sanger sequencing and RNA-seq also showed that, skipping of exon 14 and 15, and retention of intron 14 could also be observed in endogenous *RPGR* RNA in these human cells (**Figure 3C, 3D, and S3**). Taken

together, these findings demonstrated high levels of splicing diversity near exon 14 in endogenous *RPGR* expression.

The E585K mutation led to compromised RPGR splicing diversity.

The missense mutation E585K occurs one bp upstream to the canonical splice 5'ss of exon 14. Therefore, the E585K mutation was selected for further functional analysis to explore its potential impact on RNA splicing. To analyze whether and how this mutation disrupted splicing of *RPGR*, the entire genomic region spanning *RPGR* Exon 13, intron 13, exon 14, intron 14, and exon 15 were cloned into pEGFP-N3 (**Figure 4A**), and overexpressed in human 293T cells. RT-PCR showed different splicing patterns between wildtype and the c.1753G>A mutant (**Figure 4B**). Similar to endogenous *RPGR* RNA, the wildtype mini-gene was found to exhibit high splicing diversity with four distinct splicing events (**Figure 4B-E**). As confirmed by Sanger sequencing, the exon 14 skipping, E14c/E14s, and intron 14 retention accounted for 34.9%, 34.5% and 30.5% of the total wildtype splicing products respectively, while in the E585K mini-gene, E14s became predominant (86.9%) with a small fraction of transcripts containing the retended intron 14 (13.1%) (**Figure 4C, 4E-G**). Western-blot indicated that fusion in-frame EGFP protein was detected for wildtype mini-gene but not for the E585K mutant (**Figure 4D**). Human splicing finder predicted that the mutation a novel binding site for SRp55 protein (**Table S4**), which might contribute to compromised splicing diversity and enhanced usage of the alternative 5'ss.

Discussion

In the current study, our WES identified two hemizygotic *RPGR* mutation in two male RP

patients, including a novel E585K missense mutation in a patient with early onset but slow disease progression, and a frameshift deletion E998Gfs*78 in a patient with RP sine pigmento. Functional studies showed the missense mutation E585K suppressed the canonical 5'ss and skipping of E14, and enhanced an alternative 5'ss, leading to compromised splicing diversity and predomination of E14s-containing transcripts. Our results thus added to *RPGR* mutation spectrum and genotype-phenotype correlation in RP, and moreover, implicated the compromised *RPGR* splicing diversity could play a potential role in RP etiology.

RPGR undergoes complex alternative splicing in its RNA expression, and gives rise to multiple isoforms (Charng et al. 2019). Both *RPGR*_{ex1-19} and *RPGR*_{ORF15} isoforms localize to the connecting cilium of photoreceptor, and play a role of sorting and trafficking proteins from inner to outer segment (Zhang et al. 2019). However, emerging evidence indicates that *RPGR*_{ORF15} is more closely involved in RP (Talib et al.). *RPGR*_{ORF15} is highly expressed in retina, and is formed by exon 15 extending into intron 15 due to skipping of splicing, generating a repetitive sequence rich in Glutamic acid and Glycine residues and a highly conserved basic C terminus (Tee et al. 2016). Previous studies have shown that the domain interacts with WHRLN, a scaffold protein that plays a role in assembling of cytoskeleton in ear and photoreceptor (Zhang et al. 2019) (Wright et al. 2012). By far, all disease causing mutations are found to affect *RPGR*_{ORF15}, which is highly expressed in retina (Megaw et al. 2015b). Previous studies have shown that most of *RPGR* mutations result in truncated *RPGR*_{ORF15} and loss of C-terminal domain. In line with these previous findings, the two *RPGR* mutations identified in our RP patients also affected *RPGR*_{ORF15} function. As shown in **Figure 1A**, the frameshift deletion c.2993_2997del (p.E998Gfs*78) mutation in exon 15 led to a truncated

glutamic acid/Glycine-rich region and loss of the basic domain, which resulted in RP sine pigmento with high myopia in Patient 2. Before our study, there has be no RPGR mutation to be associated with RP sine pigmento. In addition, the presumed missense mutation E585K probably disrupted the splicing of normal ORF15 exon, and caused typical RP phenotypes in Patient 1. Such findings therefore added to the mutation spectrum affecting the ORF15 exon, and further underlined importance of the alternative exon in RP.

Alternatively splicing plays a role in cell differentiation, lineage determination, tissue identity acquisition and maintenance, and organ development (Baralle and Giudice 2017). A delicate splicing needs spliceosome and splicing cis-elements, including 5'ss , branch point sequence, polypyrimidine tract and splice acceptors (Ramanouskaya and Grinev 2017). Alternations of these elements could affect the efficiency, or disrupt proper splicing. In the current study, we confirmed a novel alternative 5'ss by both RT-PCR and RNA-Seq, which resulted in a 4-bp frameshift deletion and a shorter exon 14 (E14s). Interestingly, it is noted that E14s existed in multiple cell types, with a high abundance in human 293T cells (**Figure 3B**). In contrast, in RPE cells where RPGR_{ORF15} was highly expressed, the abundance of E14s was much lower, indicating the potential importance of *RPGR* splicing regulation near exon 14.

The consensus sequence of 5'ss is CAG/GUAAGUAU, in which the GU after CAG is the donor site (Ohno et al. 2018). The last base before donor site is conserved as 'G' in 80% of 5'ss (Demirci et al. 2004). In the present study, the c.1753G>A (E585K) mutation, which was predicted to introduce a missense mutation at protein level, located one bp upstream of the canonical 5'ss of exon 14, and thus might affect splicing. Our *in vitro* mini-gene assay

confirmed such possibility by showing the first evidence of a presumed benign ‘missense’ mutation in exon 14 that in fact could disrupt *RPGR* function. Previous studies showed that mutations at the last base in exons usually result in exon skipping or intron retention. For example, exon skipping could be found in around 70% mutations at the last base of exons in cancers (Jung et al. 2015). In RP, a mutation c.213G>A (G52R) has been reported to disrupt *RPGR* splicing by generating an aberrant transcript due to skipping of exon 2 (Demirci et al. 2004). However, it didn’t fit the whole story in our study. Our mini-gene assay showed that wildtype *RPGR* has a complex pattern near exon 14 that give rise to multiple splicing products in this region, including skipping of exon 14, retention of intron 14, splicing with the canonical 5’ss, as well as with the alternative 5’ss. In RPE, splicing with the canonical 5’ss showed a high priority, which generates the ‘canonical’ functional *RPGR*. Splicing with alternative 5’ss, which might generate frameshift *RPGR* transcripts, were found to exist in multiple human cell types. These findings suggest that both skipping of exon 14 and the alternative 5’ss exist in wildtype and various normal human tissues or cells, and are unlikely to cause RP by the two events directly. Instead, E585K enhanced the alternative 5’ss, and suppressed all the other events. As a result, it led to predominant transcripts containing E14s, and thus compromised splicing diversity (**Figure 5**). Such findings together with point to the importance of delicate balance in alternative splicing regulation, which maintain *RPGR* function.

In summary, we identified two mutations in *RPGR* in RP patients with different clinical features, among which a novel missense mutation c.1753G>A (E585K) was found to cause compromised splicing diversity of *RPGR* exon 13-15 junctions in the ORF15 region, and led to predomination of alternative spliced transcripts. Our findings expand our knowledge about

the mutation spectrum in RP, and provide us a new insight into alternative splicing regulation of *RPGR* in RP.

Acknowledgements

The authors thank the patients and family members for collaborating with this study. This study was supported in part by grants from the National Natural Science Foundation of China (No. 31671311 and 81371033), the National first-class discipline program of Light Industry Technology and Engineering (LITE2018-14), the “Six Talent Peak” Plan of Jiangsu Province (No. SWYY-127), the Innovative and Entrepreneurial Talents of Jiangsu Province, the Program for High-Level Entrepreneurial and Innovative Talents Introduction of Jiangsu Province, Natural Science Foundation of Guangdong Province, Guangdong High-level Personnel of Special Support Program and Yangfan Plan of Talents Recruitment Grant, Taihu Lake Talent Plan, and Fundamental Research Funds for the Central Universities (JUSRP51712B and JUSRP1901XNC), Chinese Postdoctoral Science Foundation (No.2018M642156) and Jiangsu Planned Projects for Postdoctoral Research Funds (No.2018K040A), Youth foundation of Jiangsu Natural Science Foundation (No. BK20190599).

Contributors

YL and JC designed the project and drafted the manuscript. JR, PH, EL and JC built up the collaboration network. JW, XP, QY and PH collected patients and contributed to the diagnosis of patients. ZX and JP extracted genomic DNA and performed exome data analysis together with JC. RG, SL, JZ performed Sanger sequencing. YL and CR contributed to Mini-gene assay

and WB. JR, PH and JC reviewed study design. All of the authors revised and approved the final manuscript.

Competing interests

None declared.

Patient consent

Obtained.

References

- Baralle, F. E., and J. Giudice. 2017. "Alternative splicing as a regulator of development and tissue identity." *Nat Rev Mol Cell Biol* 18 (7):437-451. doi: 10.1038/nrm.2017.27.
- Beltran, W. A., A. V. Cideciyan, A. S. Lewin, W. W. Hauswirth, S. G. Jacobson, and G. D. Aguirre. 2014. "Gene augmentation for X-linked retinitis pigmentosa caused by mutations in RPGR." *Cold Spring Harb Perspect Med* 5 (2):a017392. doi: 10.1101/cshperspect.a017392.
- Cehajic Kapetanovic, J., M. E. McClements, C. Martinez-Fernandez de la Camara, and R. E. MacLaren. 2019. "Molecular Strategies for RPGR Gene Therapy." *Genes (Basel)* 10 (9). doi: 10.3390/genes10090674.
- Charng, J., A. V. Cideciyan, S. G. Jacobson, A. Sumaroka, S. B. Schwartz, M. Swider, A. J. Roman, R. Sheplock, M. Anand, M. C. Peden, H. Khanna, E. Heon, A. F. Wright, and A. Swaroop. 2019. "Variegated yet non-random rod and cone photoreceptor disease patterns in RPGR-ORF15-associated retinal degeneration." *Hum Mol Genet* 28 (1):175. doi: 10.1093/hmg/ddy342.
- Chen, J. H., C. Huang, B. Zhang, S. Yin, J. Liang, C. Xu, Y. Huang, L. P. Cen, T. K. Ng, C. Zheng, S. Zhang, H. Chen, C. P. Pang, and M. Zhang. 2016. "Mutations of RagA GTPase in mTORC1 Pathway Are Associated with Autosomal Dominant Cataracts." *PLoS Genet* 12 (6):e1006090. doi: 10.1371/journal.pgen.1006090.
- Demirci, F. Y., A. L. Radak, B. W. Rigatti, T. S. Mah, and M. B. Gorin. 2004. "A presumed missense mutation of RPGR causes abnormal RNA splicing with exon skipping." *Am J Ophthalmol* 138 (3):504-5. doi: 10.1016/j.ajo.2004.04.019.
- Desmet, F. O., D. Hamroun, M. Lalonde, G. Collod-Beroud, M. Claustres, and C. Beroud. 2009. "Human Splicing Finder: an online bioinformatics tool to predict splicing signals." *Nucleic Acids Res* 37 (9):e67. doi: 10.1093/nar/gkp215.
- Dobin, A., and T. R. Gingeras. 2015. "Mapping RNA-seq Reads with STAR." *Curr Protoc Bioinformatics* 51:11 14 1-11 14 19. doi: 10.1002/0471250953.bi1114s51.
- Ferrante, M. I., G. Giorgio, S. A. Feather, A. Bulfone, V. Wright, M. Ghiani, A. Selicorni, L. Gammara, F. Scolari,

- A. S. Woolf, O. Sylvie, L. Bernard, S. Malcolm, R. Winter, A. Ballabio, and B. Franco. 2001. "Identification of the gene for oral-facial-digital type I syndrome." *Am J Hum Genet* 68 (3):569-76. doi: 10.1086/318802.
- Hamel, C. 2006. "Retinitis pigmentosa." *Orphanet J Rare Dis* 1:40. doi: 10.1186/1750-1172-1-40.
- Huyghe, J. R., A. U. Jackson, M. P. Fogarty, M. L. Buchkovich, A. Stancakova, H. M. Stringham, X. Sim, L. Yang, C. Fuchsberger, H. Cederberg, P. S. Chines, T. M. Teslovich, J. M. Romm, H. Ling, I. McMullen, R. Ingersoll, E. W. Pugh, K. F. Doheny, B. M. Neale, M. J. Daly, J. Kuusisto, L. J. Scott, H. M. Kang, F. S. Collins, G. R. Abecasis, R. M. Watanabe, M. Boehnke, M. Laakso, and K. L. Mohlke. 2013. "Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion." *Nat Genet* 45 (2):197-201. doi: 10.1038/ng.2507.
- Jung, Hyunchul, Donghoon Lee, Jongkeun Lee, Donghyun Park, Yeon Jeong Kim, Woong-Yang Park, Dongwan Hong, Peter J. Park, and Eunjung Lee. 2015. "Intron retention is a widespread mechanism of tumor-suppressor inactivation." *Nature Genetics* 47 (11):1242-1248. doi: 10.1038/ng.3414.
- Khanna, H. 2018. "More Than Meets the Eye: Current Understanding of RPGR Function." *Adv Exp Med Biol* 1074:521-538. doi: 10.1007/978-3-319-75402-4_64.
- Kumar, P., S. Henikoff, and P. C. Ng. 2009. "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm." *Nat Protoc* 4 (7):1073-81. doi: 10.1038/nprot.2009.86.
- Li, H., and R. Durbin. 2009. "Fast and accurate short read alignment with Burrows-Wheeler transform." *Bioinformatics* 25 (14):1754-60. doi: 10.1093/bioinformatics/btp324.
- Liu, M. M., and D. J. Zack. 2013. "Alternative splicing and retinal degeneration." *Clin Genet* 84 (2):142-9. doi: 10.1111/cge.12181.
- Lyraki, R., R. Megaw, and T. Hurd. 2016. "Disease mechanisms of X-linked retinitis pigmentosa due to RP2 and RPGR mutations." *Biochem Soc Trans* 44 (5):1235-1244. doi: 10.1042/BST20160148.
- Megaw, R. D., D. C. Soares, and A. F. Wright. 2015a. "RPGR: Its role in photoreceptor physiology, human disease, and future therapies." *Exp Eye Res* 138:32-41. doi: 10.1016/j.exer.2015.06.007.
- Megaw, Roly D., Dinesh C. Soares, and Alan F. Wright. 2015b. *RPGR: Its role in photoreceptor physiology, human disease, and future therapies*.
- Meindl, A., K. Dry, K. Herrmann, F. Manson, A. Ciccodicola, A. Edgar, M. R. Carvalho, H. Achatz, H. Hellebrand, A. Lennon, C. Migliaccio, K. Porter, E. Zrenner, A. Bird, M. Jay, B. Lorenz, B. Wittwer, M. D'Urso, T. Meitinger, and A. Wright. 1996. "A gene (RPGR) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3)." *Nat Genet* 13 (1):35-42. doi: 10.1038/ng0596-35.
- Ohno, K., J. I. Takeda, and A. Masuda. 2018. "Rules and tools to predict the splicing effects of exonic and intronic mutations." *Wiley Interdiscip Rev RNA* 9 (1). doi: 10.1002/wrna.1451.
- Parmeggiani, F., V. Barbaro, A. Migliorati, P. Raffa, P. Nespeca, K. De Nadai, C. Del Vecchio, G. Palu, C. Parolin, and E. Di Iorio. 2017. "Novel variants of RPGR in X-linked retinitis pigmentosa families and genotype-phenotype correlation." *Eur J Ophthalmol* 27 (2):240-248. doi: 10.5301/ejo.5000879.
- Patnaik, S. R., R. K. Raghupathy, X. Zhang, D. Mansfield, and X. Shu. 2015. "The Role of RPGR and Its Interacting Proteins in Ciliopathies." *J Ophthalmol* 2015:414781. doi: 10.1155/2015/414781.
- Ramanouskaya, T. V., and V. V. Grinev. 2017. "The determinants of alternative RNA splicing in human cells." *Mol Genet Genomics* 292 (6):1175-1195. doi: 10.1007/s00438-017-1350-0.
- Ramensky, V., P. Bork, and S. Sunyaev. 2002. "Human non-synonymous SNPs: server and survey." *Nucleic*

- Acids Res* 30 (17):3894-900.
- Ratnapriya, R., O. A. Sosina, M. R. Starostik, M. Kwicklis, R. J. Kapphahn, L. G. Fritsche, A. Walton, M. Arvanitis, L. Gieser, A. Pietraszkiewicz, S. R. Montezuma, E. Y. Chew, A. Battle, G. R. Abecasis, D. A. Ferrington, N. Chatterjee, and A. Swaroop. 2019. "Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration." *Nat Genet* 51 (4):606-610. doi: 10.1038/s41588-019-0351-9.
- Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J. P. Mesirov. 2011. "Integrative genomics viewer." *Nat Biotechnol* 29 (1):24-6. doi: 10.1038/nbt.1754.
- Schwahn, U., S. Lenzner, J. Dong, S. Feil, B. Hinzmann, G. van Duijnhoven, R. Kirschner, M. Hemberger, A. A. Bergen, T. Rosenberg, A. J. Pinckers, R. Fundele, A. Rosenthal, F. P. Cremers, H. H. Ropers, and W. Berger. 1998. "Positional cloning of the gene for X-linked retinitis pigmentosa 2." *Nat Genet* 19 (4):327-32. doi: 10.1038/1214.
- Talib, M., M. J. van Schooneveld, C. Van Cauwenbergh, J. Wijnholds, J. B. Ten Brink, R. J. Florijn, N. E. Schalijs-Delfos, G. Dagnelie, M. M. van Genderen, E. De Baere, M. A. Meester-Smoor, J. De Zaeytijd, F. P. M. Cremers, L. I. van den Born, A. A. Thiadens, C. B. Hoyng, C. C. Klaver, B. P. Leroy, A. A. Bergen, and C. J. F. Boon. 2018. "The Spectrum of Structural and Functional Abnormalities in Female Carriers of Pathogenic Variants in the RPGR Gene." *Invest Ophthalmol Vis Sci* 59 (10):4123-4133. doi: 10.1167/iovs.17-23453.
- Tee, J. J. L., Y. Yang, A. Kalitzeos, A. Webster, J. Bainbridge, R. G. Weleber, and M. Michaelides. 2018. "Characterization of Visual Function, Interocular Variability and Progression Using Static Perimetry-Derived Metrics in RPGR-Associated Retinopathy." *Invest Ophthalmol Vis Sci* 59 (6):2422-2436. doi: 10.1167/iovs.17-23739.
- Tee, J. J., A. J. Smith, A. J. Hardcastle, and M. Michaelides. 2016. "RPGR-associated retinopathy: clinical features, molecular genetics, animal models and therapeutic options." *Br J Ophthalmol* 100 (8):1022-7. doi: 10.1136/bjophthalmol-2015-307698.
- Tsang, S. H., and T. Sharma. 2018. "Retinitis Pigmentosa (Non-syndromic)." *Adv Exp Med Biol* 1085:125-130. doi: 10.1007/978-3-319-95046-4_25.
- Vervoort, R., A. Lennon, A. C. Bird, B. Tulloch, R. Axton, M. G. Miano, A. Meindl, T. Meitinger, A. Ciccodicola, and A. F. Wright. 2000. "Mutational hot spot within a new RPGR exon in X-linked retinitis pigmentosa." *Nat Genet* 25 (4):462-6. doi: 10.1038/78182.
- Wang, K., M. Li, and H. Hakonarson. 2010. "ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data." *Nucleic Acids Res* 38 (16):e164. doi: 10.1093/nar/gkq603.
- Weckx, S., J. Del-Favero, R. Rademakers, L. Claes, M. Cruts, P. De Jonghe, C. Van Broeckhoven, and P. De Rijk. 2005. "novoSNP, a novel computational tool for sequence variation discovery." *Genome Res* 15 (3):436-42. doi: 10.1101/gr.2754005.
- Wright, R. N., D. H. Hong, and B. Perkins. 2011. "Misexpression of the constitutive Rpgre(ex1-19) variant leads to severe photoreceptor degeneration." *Invest Ophthalmol Vis Sci* 52 (8):5189-201. doi: 10.1167/iovs.11-7470.
- Wright, R. N., D.-H. Hong, and B. Perkins. 2012. *RpgreORF15 Connects to the Usher Protein Network through Direct Interactions with Multiple Whirlin Isoforms*.
- Zhang, Q., J. C. Giacalone, C. Searby, E. M. Stone, B. A. Tucker, and V. C. Sheffield. 2019. "Disruption of RPGR protein interaction network is the common feature of RPGR missense variations that cause XLRP." *Proc Natl Acad Sci U S A* 116 (4):1353-1360. doi: 10.1073/pnas.1817639116.

443 **Table 1. Characteristics of *RPGR* mutations resulting from filtering of exome sequencing data.**

444 NF, not found; NA, not available

No.	Chromosomal position ^a	Gene symbol	mRNA	cDNA change (Protein change)	Functional prediction		dbSNP	Public /Control exomes	Sanger Sequencing
					SIFT	Polyphen			
1	chrX:38147114	<i>RPGR</i>	NM_001034853	c.1753G>A ^d (E585K)	tolerated	Benign	NF	NF	Confirmed
2	chrX:38145255	<i>RPGR</i>	NM_001034853	c.2993_2997delAAGGG (E998Gfs*78)	NA	NA	NF	NF	Confirmed

445 ^aChromosomal positions are given according to the UCSC hg19 reference assembly.

446 ^bThe mutation is at the last base of exon 14, 1 bp upstream of the canonical 5'ss.

447

448

449

450 **Table 2. Clinical features of RP patients with *RPGR* mutations in the current study.**

Patient (Mutation)	Sex	Age at recruitment	Age of onset	Visual acuity		Details of RP phenotype	Other ocular symptoms
				OD	OS		
Patient 1 (c.1753G>A ^a)	Male	50	10	LP	LP	Bilateral; night blindness; severe, diffuse, granular appearance and atrophy of RPE; bone-spicule pigmentation; arteriolar attenuation in mid-peripheral retina	Posterior subcapsular cataracts
Patient 2 (c.2993_2997del)	Male	20	7-8	HM	0.02	Bilateral; night blindness; mild bone-spicule pigmentation, and unrecordable ERG	High myopia

451 NA: not available

452 HM, hand motion; LP, light perception

453 RPE, retinal pigment epithelium; ERG, electroretinogram

454 ^aThis variant is at the last base of exon 14, 1bp upstream to the donor site, and thus might had potential impact on RNA splicing.

455

456 Figure Legend

457 **Figure 1.** Two novel *RPGR* mutations identified by WES in RP patients. (A). Location of the
458 two mutations on *RPGR*_{L-19} and/or *RPGR*_{ORF15}. The Regulator of Chromosome Condensation
459 (RCC1) like domains on both isoforms, the Glutamic acid/Glycine-rich domain and Basic
460 domain (BD) are shown as well. (B) and (C). Alignment of exome data followed by Sanger
461 sequencing reveals two novel *RPGR* mutations in two male Chinese RP patients. (D).
462 Alignment of amino acids reveals low conservation of the glutamate residue mutated in E585K
463 among mammals. (E) Alignment of genomic sequence shows the last base of human *RPGR*
464 exon 14 is not conserved among mammals.

465 **Figure 2.** Clinical examination of patient 1 with the *RPGR* c.1753G>A mutation. (A) Fundus
466 photos of Patient 1, showing bone-spicule pigmentation. Moderate senile cataract is also
467 observed in the patient. (B) Optical coherence tomography of patient 1 shows mild inner retinal
468 layer thinning and severe outer retinal layer thinning.

469 **Figure 3.** Existing of the alternative exon 14 5'ss and exon 14 skipping in endogenous *RPGR*
470 RNA in ARPE-19 and 293T cells. (A) RT-PCR of ARPE-19 and 293T cells shows multiple
471 endogenous *RPGR* splicing events. (B) Sanger sequence of endogenous *RPGR* RNA in ARPE-
472 19 and 293T cells shows existence of both the canonical and the novel alternative 5'ss in *RPGR*
473 E14. Sanger sequencing confirmed skipping of exon 14 and 15 in endogenous *RPGR* RNA in
474 ARPE-19 (C) and 293T cells (D).

475 **Figure 4.** Compromised *RPGR* splicing diversity in the c.1753G>A mutant. (A) Construction
476 of mini-gene to investigate the effects of c.1753G>A on splicing. (B) RT-PCR of 293T cells

transfected with wildtype or c.1753G>A mini-genes show distinct *RPGR* splicing patterns. (C) Quantification of the RT-PCR results shown in (B). (D) Western blot of 293T cells transfected with EGFP, wildtype or c.1753G>A mini-gene. The c.1753G>A mutation suppresses the canonical 5'ss in E14, resulting in undetectable fusion GFP protein with canonical RPGR. (E) Sanger sequencing confirms existence of both the canonical and the novel alternative 5'ss in *RPGR* exon 14 in wildtype. (F) Sanger sequencing confirms skipping of exon 14 in wildtype, which is suppressed in the c.1753G>A mutant. (G) Sanger sequencing confirms retention of intron 14 in both wildtype and the c.1753G>A mutant.

Figure 5. A schematic diagram illustrating compromised *RPGR* splicing diversity due to the c.1753G>A mutation. (A) For the wildtype *RPGR*, the canonical 5'ss (red) generates a canonical *RPGR* with exon 13, 14 and 15. The alternative 5'ss (green) generates a transcript with the last 4bp of exon 14 deleted (E14s), resulting in a frameshift *RPGR*. Skipping of exon14 generates a frameshift transcript, resulting in nonsense mediated mRNA decay. And retention of intron 14 produces an in-frame but longer transcript. (B) Splicing events of the canonical 5'ss and skipping of exon14 are suppressed due to the c.1753G>A mutation (purple), giving rise to predominant transcripts containing E14s with a small fraction of transcripts containing the retended intron 14, and thus compromised splicing diversity.

A **B**

A **B** **C** **D**

A **B** **C** **D** **E** **F** **G**

A

Wildtype

Splicing Event

1

Canonical SD

Reading frame

inframe

2

Alternative SD

frameshift

3

Exon Skipping

frameshift

4

Intron Retention

inframe

B

c.1753G>A

Splicing Event

●

Canonical SD

Reading frame

inframe

2

Alternative SD

frameshift

●

Exon Skipping

frameshift

4

Intron Retention

inframe