Identification of a novel pathogenic missense mutation in PRPF31 using whole exome sequencing: a case report

Laura Bryant,1 Olga Lozynska,1 Anson Marsh,1 Tyler E Papp,1 Lucas van Gorder,1 Leona W Serrano,2 Xiaowu Gai,3,4 Albert M Maguire,2 Tomas S Aleman,2 Jean Bennett1

ABSTRACT

Background Variants in PRPF31, which encodes pre-mRNA processing factor 31 homolog, are known to cause autosomal-dominant retinitis pigmentosa (adRP) with incomplete penetrance. However, the majority of mutations cause null alleles, with only two proven pathogenic missense mutations. We identified a novel missense mutation in PRPF31 in a family with adRP.

Methods We performed whole exome sequencing to identify possible pathogenic mutations in the proband of a family with adRP. Available affected family members had a full ophthalmological evaluation including kinetic and two-colour dark adapted static perimetry, electrophotography and multimodal imaging of the retina. Two patients had evaluations covering nearly 20 years. We carried out segregation analysis of the probable mutation, PRPF31 c.590T>C. We evaluated the cellular localisation of the PRPF31 variant (p.Leu197Pro) compared with the wildtype PRPF31 protein.

Results PRPF31 c.590T>C segregated with the disease in this four-generation autosomal dominant pedigree. There was intrafamilial variability in disease severity. Nystagmus and mid-peripheral scotomas presented from the second to the fourth decade of life. There was severe rod > cone dysfunction. Visual acuity (VA) was relatively intact and was maintained until later in life, although with marked interocular asymmetries. Laboratory studies showed that the mutant PRPF31 protein (p.Leu197Pro) does not localise to the nucleus, unlike the wildtype PRPF31 protein. Instead, mutant protein resulted in punctate localisation to the cytoplasm.

Conclusions c.590T>C is a novel pathogenic variant in PRPF31 causing adRP with incomplete penetrance. Disease may be due to protein misfolding and associated abnormal protein trafficking to the nucleus.

BACKGROUND

Retinitis pigmentosa (RP) is the most common form of inherited retinal degeneration, affecting 1 in 2500–7000 people1 Over 60 genes have been shown to cause RP (https://sph.uth.tmc.edu/retnet/home.htm). It can be inherited in an autosomal recessive, autosomal dominant or x linked inheritance pattern, with different genes associated with each inheritance pattern.1 Approximately 30%–40% of the cases of retinitis pigmentosa are autosomal dominant (adRP).2 Dominant disease can be due to either a toxic gain of function in which the mutation causes the protein to have a directly toxic effect on the cell, or haploinsufficiency, where half the normal amount of protein is insufficient to fulfil the needs of the cell.

PRPF31 encodes a ubiquitously expressed splicing factor.3 It links the U4/U6 complex with U5, creating the tri-snRNP of the spliceosome.4 PRPF31 mutations are known to cause adRP with incomplete penetrance (RP11, OMIM 600138).5-10 It is interesting that PRPF31 mutations result in a retina specific phenotype when the splicing factor is ubiquitously expressed. A retina specific isoform has not been identified that can explain the tissue specific susceptibility of the retina to a heterogeneous mutation in PRPF31.11 The retina appears to simply have a higher dependence on this splicing factor than other tissues. Deery et al speculated this could be due to the need to constantly replenish disc proteins in the outer segments, resulting in a higher splicing load than in other cell types.12 It is also possible that there are splicing factors in other cell types, but not the retina, that can compensate for the mutant PRPF31.

One explanation for the incomplete penetrance seen in families with PRPF31 mutations is the variable expression levels of PRPF31.8 There is a critical level of PRPF31 needed to avoid retinal degeneration. If both alleles are wildtype, the critical level is exceeded and retinal degeneration is avoided. However, if the expression level of one wildtype allele is high enough, a carrier of a pathogenic allele will still reach the critical level of wildtype protein and be asymptomatic. If the wildtype allele has an average or low level of expression, a carrier of a pathogenic allele will develop retinal degeneration.

We identified a novel missense mutation in PRPF31. Most mutations in PRPF31 are truncations, deletions or frameshift mutations which result in a null allele.13 However, two missense mutations have been shown to be pathogenic12 and are located within 20 amino acids of the novel mutation seen in the family in this study, suggesting it could be in an important domain for protein function. We analysed the segregation of the mutation within the family and the impact of the mutation on localisation of the protein and concluded that it was the disease-causing mutation in this family. The results increase our understanding of a potential pathogenetic mechanism of PRPF31-mediated disease, thereby paving the way for development of a treatment.
METHODS
Charactrisation of the clinical phenotype: Informed consent was obtained after explanation of the nature of the study; procedures complied with the Declaration of Helsinki and were approved by the institutional review board (IRB; University of Pennsylvania IRB #808828, 815348). Patients had a comprehensive eye examination, best-corrected visual acuity (VA), Goldmann kinetic visual fields and two-colour dark adapted perimetry using a modified Humphrey visual field. Spectral-domain optical coherence tomography (SD-OCT) imaging was performed with a Cirrus HD-OCT (Carl Zeiss Meditec AG, Dublin, USA) instrument and 6 mm horizontal sections crossing the anatomic fovea. Electrotoretinography (ERG) was performed with a computer-based system (EPIC-XL, LKC Technologies, Gaithersburg, Maryland, USA) following ISCEV standards. Whole exome sequencing: Testing was carried out on DNA samples from human subjects after obtaining written informed consent on an Institutional Review Board (IRB) approved protocol (#808828). A sample from the proband was screened previously for mutations in rhodopsin (RHO), peripherin/RDS and ROM1 (Carver Lab, University of Iowa, Iowa City, Iowa, USA, 1995) and found to be negative. For whole exome sequencing, we performed target enrichment using Agilent SureSelect target enrichment system and whole exome sequencing was performed on Illumina HiSeq2000 at the Penn Genome Frontiers Institute (PGFI). BWA (V0.5.9-r16) was used to align the sequence reads to the human reference genome GRCh37. Samtools (V0.1.12 or r859) was used to make initial SNP and indel calls which were then further refined and filtered with a custom programme that used the Benjamini and Hochberg correction based on quality values of all bases to minimise false positive calls. A coverage depth cut-off of 10× was applied. Resulting variant calls were annotated using our custom human bp codon database. Additional annotations of each variant call were provided using data sets downloaded from the 1000 Genomes Project website (http://www.1000genomes.org/), the NHLBI Exome Sequencing Project Exome Variant Server (http://evs.gs.washington.edu/EVS/) and the UCSC Genome Browser (http://genome.ucsc.edu/). These annotations included allele frequencies, SIFT and PolyPhen predictions and phastCons conservation scores. Custom scripts were also developed and used to identify candidate variants that fit different filtering criteria, such as genetic models. We focused on novel variants at the time of the study, limiting the analysis to variants that was not reported by EVS at the time and only present in the proband.

This yielded a short list of genes which were prioritised based on mutation type and likelihood to cause adRP. These include a list of 105 novel candidate variants from 104 unique genes (online supplementary table 1). The PRPF31 variant was the strongest candidate. The proband was negative for mutations in the CCR4-NOT transcription complex, subunit 3 (CNOT3) gene which has been postulated as one modifier gene responsible for the incomplete penetrance of PRPF31-adRP.

Segregation analysis: We used PCR to amplify the DNA region that included the variant from the proband and all relatives who provided DNA using Phusion (New England Biolabs, Ipswich, Massachusetts, USA) and the following primers synthesised by Invitrogen (Carlsbad, CA): GAGCCCTCTGAGTTCCCG and GCCAAAGGCCCCATTCTAC. The PCR product was sent for Sanger sequencing at the Penn Genomics Analysis Core and visualised using SnapGene software (from GSL Biotech (Chicago, Illinois, USA); available at snapgene.com).

Cloning: We cloned the sequences for PRPF31 from a cDNA library generated from 293 T cells. We used Q5 polymerase (New England Biolabs) and custom primers synthesised by Invitrogen (see online supplementary table 1). The PCR product was TOPO cloned and the sequences were verified by Sanger sequencing at the Penn Genomics Analysis Core. The Refseq sequences used for PRPF31 areNM_015629 and NP_056444. The coding sequences were cloned into an expression vector with a C-terminal HA tag using In-Fusion (Takara Bio, Mountain View, California, USA). The expression vector used a chicken β actin (CBA) promoter with a cytomegalovirus enhancer promoter to drive expression. We used the same vector, except replacing PRPF31 with the cDNA encoding enhanced Green Fluorescent Protein (eGFP); but no HA tag) as a control for transfection efficiency.

Mutagenesis: We used site directed mutagenesis (Quik-Change II Site-Directed Mutagenesis Kit from Agilent Technologies; Santa Clara, California, USA) to introduce the p.L197P variant into the cloned PRPF31 using the following primers: GATGCGGTGCTGGAGCGCTGCTCCAGCGCGCAT GTCCGCAG and CTGCCGACATGGCGCTGAGCCGAACGCC TCCAAGCACCGCATC.

Transfection: ARPE19 cells were maintained in DMEM F12 media (Invitrogen) supplemented with 10% fetal bovine serum. We plated the cells in four well chamber slides. The cells were transfected using Lipofectamine LTX with plus reagent (Thermo Fisher Scientific; Agawam, Massachusetts, USA) and fixed 48 hours post-transfection.

Immunofluorescent staining: The ARPE19 cells were fixed for 15 min in 4% paraformaldehyde. They were rinsed three times with Dulbecco’s phosphate-buffered saline (Corning 21-030-CV). We then blocked the cells for 1 hour using a blocking buffer consisting of 10% normal goat serum and 0.5% Triton X-100 in PBS. The slides were then incubated overnight at four degrees with an antibody against the HA tag (Cell Signalling Technology #3724; Danvers, Massachusetts, USA) at a 1:800 dilution. The slides were then incubated at room temperature for 3 hours in Alexafluor 488 Goat anti-Rabbit (Thermo Fisher Scientific A-11034). The cells were mounted using Fluoromount-G mounting media contained 4',6-diamidino-2-phenylindole. The cells were imaged using an Olympus FV1000 confocal microscope with a 60× oil immersion objective. The transfections were repeated three times.

RESULTS
The proband (II-3) presented at age 37 with a 5-year history of nyctalopia. Snellen VA was 20/25 and 20/30, for the right and left eye, respectively, with a mild myopic refractive error (table 1). Kinetic visual fields measured with a V-4e target showed an absolute scotoma superiorly and inferiorly from the blind spot (figure 1). Her rod ERGs to a dim flash were non-detectable, had reduced a-wave and b-wave amplitudes in response to a standard flash with a negative configuration waveform and showed cone-mediated responses that were only mildly reduced in amplitude (online supplementary figure 1). The ERG supports an inner retinal dysfunction, perhaps as a consequence of inner retinal remodelling, and is consistent with previous observations in RP and PRPF31-adRP. The proband’s father in his early 60 s (I-1), exemplified intrafamilial variability of severity (table 1). When examined years ago, he showed minimal pigmentary changes and was visually asymptomatic; he had a sister with confirmed retinal degeneration and a visually impaired brother who had not been formally diagnosed as a retinal degeneration (figure 1). Two of the proband’s siblings...
(II-1 and II-2) and a first cousin had also become symptomatic in their early 30s with nystagmus and difficulties navigating obstacles and had been diagnosed with RP; her affected siblings (II-1, II-2) were evaluated by us at least once (figure 1A; table 1).

When examined near presentation (II-1; II-3) patients had VAs better than 20/30 and showed a mid-peripheral pigmentary retinopathy. On longitudinal follow-up, both patients lost VA to a different extent with marked interocular asymmetries (figure 1B). Cystoid macular oedema responsive to oral acetazolamide caused temporary VA losses in II-3 at age 45 (figure 1B). Longitudinal follow-up for nearly 20 years in II-1 and II-3 demonstrated progression of the field changes and provided additional evidence of intrainfamilial variability in severity. Patient II-3 had normal peripheral visual field extent to a V4e target and an incomplete near mid-peripheral scotoma at age 47 compared with the severe visual field loss documented in her brother (II-1) at age 46 (figure 1C) who had a small island near fixation and a remnant of vision in the temporal field (figure 1C). He had only detectable cone-mediated responses that were markedly reduced in amplitude (100% of normal). In the last visit, II-3 showed a small island of vision near fixation separated from a temporal island of vision by a complete mid-peripheral scotoma; II-1 had lost his peripheral island and showed detectable vision only near fixation (figure 1C). The pattern of visual field loss in II-3 in her last visit was reminiscent to that of her brother II-1 at the earliest visit supporting similar progression to a common phenotype but variation in age of onset and severity. II-2 was examined by us only once at age 66 when he showed VA of 20/30 and 20/60 for the right and left eye, respectively, and severely constricted visual fields (figure 1C). At the last examination, the central retinal appearance of II-3 was relatively benign whereas her brothers showed a mid-peripheral pigmentary retinopathy with areas of retinal pigmented epithelium (RPE) atrophy encroaching on central islands of relative RPE preservation and a pigmentary retinopathy (figure 1E).

Searching for a molecular cause: whole exome sequencing

The family history and phenotype supported the diagnosis of adRP. Analysis of the whole exome sequencing data shows that the proband has a novel c.590T>C (p.Leu197Pro) missense variant in PRPF31 (figure 2A). This variant is not listed in dbSNP or ExAC. Two missense mutations in PRPF31 that are known to be pathogenic are p.Ala194Glu and p.Ala216Pro.9,12,26 The proximity of the new variant with these other amino acid changes suggests that this region is an important domain for protein function.12 Since PRPF31 is a splicing factor, mislocalisation from the nucleus would essentially constitute a loss of function mutation. Both the p.Ala194Glu and p.Leu197Pro variants are located in the second coiled-coil domain while the p.Ala216Pro links the coiled-coil to the (Nop) domain which is a ribonucleolar protein recognition motif.27

Four of the proband’s siblings (two affected, two unaffected) were tested for the c.590T>C mutation in PRPF31 (see figure 2A). Two affected brothers (II-1 and II-2) were heterozygous for the variant while one of the unaffected siblings (II-4) was homozygous for the wildtype allele. DNA extraction failed in the saliva sample derived from the additional unaffected brother (II-6). One of the asymptomatic, presumably unaffected siblings tested (sister, II-5) is a carrier for the c.590T>C variant. In sum, the inheritance pattern is consistent with AD disease with incomplete penetrance typically seen in PRPF31 mutations. Although the asymptomatic carrier may continue to enjoy good vision, any children that she may have who inherit the mutation will be at risk for retinal degeneration.

Since two of the previously known PRPF31 missense variants mislocalise to the cytoplasm,12 we tested the cellular localisation of the p.Leu197Pro variant. We cloned the cDNA sequence for PRPF31 from 293 T cells and verified that it matched the reference sequence. We then performed site directed mutagenesis to add the c.590T>C variant. We cloned both PRPF31WT and PRPF31L197P into expression vectors, each with driven by a constitutive promoter (the CBA promoter and cytomegalovirus enhancer (CMV/CBAj) and the transgene was tagged with the human influenza haemagglutinin (HA) marker. Another set of cells was transfected with pCBA.eGFP as a control for transfection efficiency. We transfected ARPE19 cells and analysed protein localisation 48 hours post-transfection with microscopy. Although equal numbers of cells were plated, there were more cells in the control and PRPF31WT-treated wells at the 48 hours

---

**Table 1** Clinical characteristic of the patients with (heterozygous) PRPF31L197P

<table>
<thead>
<tr>
<th>Pt #</th>
<th>Age*</th>
<th>Gender</th>
<th>Visual acuity†</th>
<th>Refraction‡</th>
<th>Kinetic perimetry§</th>
<th>Foveal thickness (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>OD</td>
<td>OS</td>
<td>OD</td>
<td>OS</td>
</tr>
<tr>
<td>II-1</td>
<td>64/M</td>
<td>LP</td>
<td>20/60</td>
<td></td>
<td>Pseudophakia</td>
<td>Pseudophakia</td>
</tr>
<tr>
<td>II-2</td>
<td>66/M</td>
<td>20/60</td>
<td>20/30</td>
<td></td>
<td>Pseudophakia</td>
<td>Pseudophakia</td>
</tr>
<tr>
<td>II-3</td>
<td>59/F</td>
<td>20/30</td>
<td>20/125</td>
<td>+0.25</td>
<td>+0.50</td>
<td>10+inf</td>
</tr>
</tbody>
</table>

All patients showed waxy pallor of the nerves with peripapillary atrophy, central islands of relative RPE preservation and a pigmentary retinopathy (figure 1).

*Age in years.
†LP, light perception.
‡Spherical equivalent.
§Visual field extent in degrees (Goldmann V-4e target); +inf denotes inferior island of vision separated from the central residual island.
OD, Ocula Dextra (Right Eye); OS, Oculi Sinistra (Left Eye).
RPE, retinal pigmented epithelium.

---

**Clinical science**
timepoint than PRPF31<sup>L197P</sup> cells, suggesting that there was some toxicity due to the mutant PRPF31 protein. Transfection efficiency, as assessed by the eGFP control, ranged from 5% to 15% from experiment to experiment.

In the PRPF31-treated cells, HA-tagged PRPF31<sup>WT</sup> protein localised exclusively to the nucleus in more than a third of the transfected cells (with protein more diffusely spread across the cytoplasm in the majority of the remaining transfected cells (figure 2B, online supplementary figure 2). In contrast, the HA-tagged PRPF31<sup>L197P</sup> protein was predominantly in a punctate pattern in the cytoplasm (figure 2B), possibly indicating that it is being targeted for degradation (see also online supplementary figure 2). This speckled pattern was nearly absent in cells transfected with PRPF31<sup>WT</sup>.

**DISCUSSION**

Mutations in PRPF31 are well known to cause adRP with incomplete penetrance. Most of these mutations are truncations, frame-shifts, splicing mutations or large deletions, all of which cause null alleles. Pathogenic missense mutations in PRPF31 are much rarer. The universal protein resource site (uniprot.org) only lists two missense mutations associated with RP11. The single nucleotide polymorphism database (DbSNP) only classifies two missense variants in PRPF31 as ‘pathogenic’, while listing three...
Figure 2  Segregation and functional analysis of the novel PRPF31 variant. (A) Sanger sequencing results for the PRPF31 c.590T>C variant. The variant was confirmed in the proband (II-3) and an affected brother (II-1). One of the unaffected sisters (II-5) is homozygous for the wildtype allele while the second unaffected sister is a carrier for the variant (II-4). An unaffected brother (II-6) could not be tested due to failed DNA purification. (B) Localisation of PRPF31 WT and PRPF31 L197P in ARPE19 cells as measured by immunofluorescence for the HA tag (green). Nuclei appear blue due to staining with DAPI. PRPF31 L197P predominantly mislocalised to the cytoplasm providing a punctate localisation pattern, while PRPF31 WT generally localised primarily to the nucleus. DAPI, 4',6-diamidino-2-phenylindole; HA, haemagglutinin.

more as ‘likely pathogenic’. Given that the RP11 phenotype involves incomplete penetrance and pathogenic missense variants are rare, functional testing is more important than usual to establish pathogenicity. The fact that the variant PRPF31 is not localised efficiently to the nucleus makes it impossible for PRPF31 L197P to be functional as a splicing factor.
The mislocalisation seen in the PRPF31 p.Leu197Pro protein variant as well as the previously described p.Ala194Glu and p.Ala216Pro variants is interesting considering they are unlikely to disrupt a nuclear localisation signal. 194 and Leu197 are located in a coiled coil domain while Ala216 is in a linker region between the coiled coil and the NOP domain.25 It seems likely that these missense mutations are destabilising the protein structure or disrupting protein folding and causing the protein to be targeted for degradation. This is consistent with the staining pattern we observed for PRPF31 L197P.

Clinically, nyctalopia, a mid-peripheral pigmented retinopathy, retina-wide rod greater than cone dysfunction and incomplete or non-penetration are hallmarks of PRPF31-adRP (RP11) and were all documented in the family described herein.10 13 26 28 The non-penetration of PRPF31-adRP has rightfully attracted most of the attention for nearly five decades. However, it is worth noting the disease’s consistent patterns. For example, relatively preserved VA into late stages of the disease is a relatively common finding, which is reassuring to patients diagnosed with this molecular subtype of adRP and consistent with the relatively preserved central rod and cone function documented in this and other reports.11 23 24 Reassurance, however, should be conveyed with caution given the variability in severity associated with this genotype. Another frequent finding is the presence of perifoveal changes and macular atrophy often adopting a bull’s eye pattern.13 28 Movement of the transition zones of degeneration from the perifovea towards the foveal centre underlie the asymmetric decline in VA witnessed in our patients. Despite the generalised severe rod dysfunction reported in this condition by ERG, this study documented substantial rod function accompanying near normal cone vision for central retinal locations late in the course of the disease confirming previous observations,10 25 Observation of such a pattern should prompt consideration of PRPF31 as a possible cause, is consistent with a mechanism of disease that is not the direct cause of the sensitivity losses but rather indirectly through photoreceptor loss and/or structural abnormalities or loss of the photoreceptor outer segment and adds PRPF31 to a growing list of genes associated with retinal degenerations where some retinal regions can show normal structure and/or rod and/or cone photoreceptor function.10 14 25 Understanding the mechanisms that confer protection from degeneration to certain individuals or retinal regions in PRPF31-adRP will lead to a better understanding of the pathophysiology of the larger group of inherited retinal degenerations as a premise to finding treatments.3

CONCLUSION

We conclude that the c.590T>C missense variant in PRPF31 is a pathogenic mutation causing adRP with incomplete penetrance. The phenotype is consistent with the phenotype seen from other PRPF31 mutations, the mutation segregates with disease in the family with incomplete penetrance and the missense mutation causes mislocalisation of the protein in vitro.

Author affiliations
1Center for Advanced Retinal and Ocular Therapeutics (CAROT) and F.M. Kirby Center for Molecular Ophthalmology, Scheie Eye Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
2Department of Ophthalmology and CAROT, Scheie Eye Institute, University of Pennsylvania Perelman School of Medicine, Perelman Center for Advanced Medicine, Philadelphia, Pennsylvania, USA
3Center for Personalized Medicine, Department of Pathology and Laboratory Medicine, Children’s Hospital Los Angeles, Los Angeles, California, USA
4Keck School of Medicine, University of Southern California, Los Angeles, California, USA

Acknowledgements We are grateful to the patients who were research volunteers and to technical assistance from Nicholas Phelps, Angela Luo and Ayodele Maja.

Contributors TSA, LS and AMM carried out patient assessments and referred potential patients to the study. LB, DL, LvG, AM, TEP, KG and JB contributed to data collection and analyses. LB, TSA and JB wrote the manuscript. All authors reviewed the manuscript.

Funding This study was funded by NIH Vision Training Grant ST32EY007035-37, a Center grant (RC-CL-0607-0389-UPA01) from Foundation Fighting Blindness to the CHOP-Penn Pediatric Center for Retinal Degenerations, the Brenda and Matthew Shapiro Stewardship and the Robert and Susan Heidenberg Investigative Research Fund for Ocular Gene Therapy, Research to Prevent Blindness, the Paul and Evania Mackall Foundation Trust, the Center for Advanced Retinal and Ocular Therapeutics, and the F.M. Kirby Foundation.

Disclaimer The senior author (JB) affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Competing interests JB is a founder of Gensight Biologics and of Limelight Bio and a scientific (non-equity-holding) founder of Spark Therapeutics. JB holds Sponsored Research Agreements (SRAs) from Biogen, Limelight Bio and REGENXBio. The coauthors report no additional conflicts.

Patient consent Obtained.

Ethics approval University of Pennsylvania Institutional Review Board (IRB).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All materials generated in this study and all data will be available to other authors. Patient-derived samples will be available under a Material Transfer Agreement.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

Clinical science