Expression of **PRPF31** and **TFPT**: regulation in health and retinal disease

Anna M. Rose¹, Amna Z. Shah¹, Naushin H. Waseem¹, Christina F. Chakarova¹, Giovanna Alfano¹, Razek G. Coussa², Radwan Ajlan², Robert K. Koenekoop² and Shomi S. Bhattacharya¹

¹Department of Genetics, UCL Institute of Ophthalmology, London EC1V 9EL, UK and ²The McGill Ocular Genetics Laboratory and Paediatric Ophthalmology, Montreal Children’s Hospital, McGill University Health Centre, Montreal H3H 1P3, Canada

Received April 26, 2012; Revised May 17, 2012; Accepted June 14, 2012

**PRPF31**, a gene located at chromosome 19q13.4, encodes the ubiquitous splicing factor PRPF31. The gene lies in a head-to-head arrangement with **TFPT**, a poorly characterized gene with a role in cellular apoptosis. Mutations in **PRPF31** have been implicated in autosomal dominant retinitis pigmentosa (adRP), a frequent and important cause of blindness worldwide. Disease associated with **PRPF31** mutations is unusual, in that there is often non-penetrance of the disease phenotype in affected families, caused by differential expression of **PRPF31**. This study aimed to characterize the basic promoter elements of **PRPF31** and **TFPT**. Luciferase reporter constructs were made, using genomic DNA from an asymptomatic individual with a heterozygous deletion of the entire putative promoter region. Fragments were tested by the dual-luciferase reporter assay in HeLa and RPE-1 cell lines. A comparison was made between the promoter regions of symptomatic and asymptomatic mutation-carrying individuals. A patient (CAN493) with adRP was identified, harbouring a regulatory region mutation; both alleles were assayed by the dual-luciferase reporter assay. Luciferase assays led to the identification of core promoters for both **PRPF31** and **TFPT**; despite their shared gene architecture, the two genes appear to be controlled by slightly different regulatory regions. One functional polymorphism was identified in the **PRPF31** promoter that increased transcriptional activation. The change was however, not consistent with the observed symptomatic–asymptomatic phenotypes in a family affected by **PRPF31**-adRP. Analysis of the mutant promoter fragment from CAN493 showed a >50% reduction in promoter activity, suggesting a disease mechanism of functional haploinsufficiency—the first report of this disease mechanism in adRP.

**INTRODUCTION**

**PRPF31** is a highly conserved gene located at chromosome 19q13.4, encoding the ubiquitous splicing factor PRPF31. The process of splicing occurs by the sequential assembly and disassembly of the spliceosomal small nuclear ribonucleoproteins (snRNPs), one crucial step being the association of the U4/U6 di-snRNP with the U5 snRNP, forming the U4/U6.U5 tri-snRNP; the assembly of the tri-snRNP is facilitated by PRPF31 protein (1–5). **PRPF31** lies in a head-to-head arrangement with **TFPT**, with the two genes partially sharing exon 1. **TFPT** was first identified as a fusion partner of E2A in some cases of paediatric pre-B cell acute lymphoblastic leukaemia (6). **TFPT** was shown to be a ubiquitously expressed gene, conserved in rodents and humans, with a high level of sequence homology, but has little similarity to other known proteins (6). Experimental work has shown that the protein plays an important role in p53-independent cellular apoptosis (7–9). There is also some evidence that it plays a role in cell cycle arrest and spermatogenesis (10,11). Mutations in **PRPF31** have been implicated in autosomal dominant retinitis pigmentosa (adRP). Retinitis pigmentosa (RP) represents a group of retinal degenerations affecting 1 in 4000 people, characterized by progressive loss of rod and cone photoreceptors, leading to significant visual handicap (12). The disease displays marked genetic heterogeneity,
with disease-causing mutations reported in over 50 genes, with all manners of Mendelian inheritance (13). Among these, mutations in six ubiquitously expressed splicing factors are known to cause adRP: PRPF3, PRPF6, PRPF8, PRPF31, SNRNP200 and RP9 (14–19). All splicing factors implicated in adRP are associated with the U4/U6 particle, with the exception of RP9, which interacts with another splicing factor, U2AF35 (20).

A key, consistent feature of PRPF31-associated adRP is the very unusual pattern of inheritance: the affected families show non-penetrance—necessitating the existence of asymptomatic mutation carriers, despite autosomal dominant inheritance. It has been suggested that non-penetrance is due to the differential expression level of PRPF31, the expression level being controlled by an allele on chromosome 19 (21). If a mutant PRPF31 allele and a ‘high-expressivity’ wild-type (WT) allele are inherited, the residual protein level is sufficient for normal retinal function. If, however, a mutant PRPF31 allele and a ‘low-expressivity’ WT allele are co-inherited, the PRPF31 level falls beneath the threshold required for normal retinal function and clinically manifest disease results; this is a variant on the phenomenon of haploinsufficiency. The notion of differential PRPF31 expression levels was confirmed when the level of PRPF31 mRNA transcript was measured by real-time quantitative polymerase chain reaction (PCR) in lymphoblastoid cell lines derived from patients and their asymptomatic relatives, using a probe designed to anneal solely to the WT allele; this demonstrated an almost 2-fold higher level of PRPF31 transcript in the asymptomatic individuals compared with the symptomatic individuals (22). The mechanism controlling the different level of expression is, as yet, unexplained, prompting this study.

As the first step in understanding the differential expression of PRPF31, it was considered necessary to define the PRPF31 core promoter. PRPF31 and TFPT are arranged as a head-to-head bidirectional gene pair, with PRPF31 on the forward strand and TFPT on the reverse strand. It was, therefore, postulated that the genes might share a bidirectional promoter. Brambillasca et al. (23) characterized the basic promoter of TFPT, leading to the definition of a fragment [FB-Luc, −759 to +104 bp relative to TFPT transcription start site (TSS)] with very weak reporter activity in both forward strand and reverse strand orientation.

It is recognized that common promoter region polymorphisms within the population are capable of significantly altering the basal level of transcription, and in some cases, this has consequences for human disease. Functional polymorphisms are able to affect transcription by changing an existing transcription factor (TF)-binding site—either increasing or decreasing TF-binding affinity or by creating a novel TF-binding site. Diseases associated with such polymorphisms are varied and include increased susceptibility to infectious disease, autoimmunity and cancer (24–26). It was, therefore, postulated that promoter polymorphisms might underlie the differential expression of PRPF31.

In this study, we identify and characterize the core promoters of two genes, PRPF31 and TFPT. We also report the functional effect of polymorphisms within the PRPF31 core promoter and investigate whether these are responsible for differential gene expression and, therefore, the clinical manifestation of PRPF31-associated adRP. Furthermore, we describe a pathogenic mutation within the PRPF31 core promoter in a family affected by adRP.

**RESULTS**

**Fragment design, amplification and cloning into pGL3**

In total, 16 fragments were assayed in the initial experiments, spanning from −2122 to +926 bp relative to the annotated PRPF31 TSS, these fragments including part of the 5′-untranslated region (5′UTR) and the upstream region of the gene (Fig. 1A). These fragments were divided into two groups: those replicating the assay of Brambillasca et al. (23) (Group 1) and new fragments (Group 2). Fragments were successfully produced by PCR using genomic DNA from a previously reported asymptomatic hemizygous individual (III.7) harbouring a heterozygous deletion of the entire putative regulatory region (27). All fragments were cloned into pGL3-basic vector in forward strand orientation (PRPF31 orientation, indicated by ‘+’), and also some in reverse orientation (TFPT orientation, indicated by ‘−’) (Supplementary Material, Tables S1 and S2). After the luciferase assay and bioinformatic analysis, three further fragments (Bi-P, Δ2-1 and Δ2-2) were amplified, cloned into pGL3-basic and assayed in both orientations (Fig. 1B).

**Dual-luciferase reporter assay to characterize the core promoter elements of PRPF31 and TFPT**

The reporter assay of the Group 1 fragments showed that fragment Δ2+ had strong promoter activity, with an ~6-fold induction over the minimal thymidine kinase promoter (pTK) [6.02 ± 0.47 (HeLa); 5.69 ± 0.29 (RPE-1)] (Fig. 2A). The remainder of the fragments showed lower activity than pTK, indicating that they are not active promoter elements. The presence of the Δ1 sequence in either orientation abolished promoter activity, demonstrating that this fragment has inhibitory activity.

The reporter assay of Group 2 fragments was first performed in the forward strand orientation, showing that no fragment had activity equivalent to pTK or stronger (Fig. 2B). Fragment P.31-Luc had near pTK activity [0.80 ± 0.15 (HeLa); 0.63 ± 0.11 (RPE-1)] and was, therefore, assessed in the reverse strand orientation, showing moderate promoter activity [2.15 ± 0.20 (HeLa); 1.30 ± 0.16 (RPE-1)].

In the light of the results, a fragment comprised of both P.31-Luc and Δ2 (Bi-P) was produced and assayed in both orientations, this showing strong promoter activity in forward strand (PRPF31) orientation [8.28 ± 0.84 (HeLa); 7.01 ± 0.68 (RPE-1)], but weak activity in the reverse strand orientation [0.63 ± 0.09 (HeLa); 0.95 ± 0.2 (RPE-1)] (Fig. 2C). Promoters controlling the expression of housekeeping genes (such as PRPF31) are usually under epigenetic control by CpG methylation and, as such, the CpG content of the Bi-P fragment was analysed, showing that the region was indeed a CpG island, supporting its likelihood of being the PRPF31 core promoter (Supplementary Material, Fig. S1).

Fragment Δ2+ was analysed bioinformatically, showing two distinct clusters of putative TF-binding sites (Supplementary Material, Fig. S2). Two deletion fragments were designed
that separated these two clusters; these fragments were also tested by the luciferase assay (Δ2-1+ and Δ2-2+, Fig. 1B). This showed that reporter activity required the presence of the whole Δ2+ fragment (Fig. 2D). The full luciferase assay data are presented in Supplementary Material, Table S3.

We concluded that Bi-P was the core promoter element controlling the expression of PRPF31, whereas P.31-Luc was the core promoter element for TFPT. Electrophoretic mobility shift assay (EMSA) was performed to validate that the Bi-P fragment was able to bind nuclear TFs. This assay showed a marked mobility shift when the Bi-P fragment was incubated with HeLa nuclear extract, this interaction being specific (Fig. 2E).

Sequence comparison between a symptomatic and an asymptomatic individual

The sequence of the entire putative promoter region was compared in the asymptomatic test individual (III.7) and his symptomatic daughter (RP15011), who also harboured the same heterozygous deletion of the region (27). Three changes were observed between these two individuals (Supplementary Material, Fig. S3). Briefly, a single-nucleotide polymorphism (SNP) was observed at rs78558230, lying at −605 relative to the PRPF31 TSS; individual III.7 carried the minor T allele, whereas RP15011 carried the C allele. Additionally, two previously unreported changes were identified that differed between the two individuals. First, a hemizygous C to G substitution at base 54617483 (hg19 co-ordinates, −1354 to TSS) was identified in individual III.7 and this base change was not present in RP15011; the change was verified as a rare polymorphism in the European population. The second unreported change was a 38 bp duplication/insertion, present at −632 to TSS, this duplication was present in III.7 but not RP15011; again, the change was verified as a normal polymorphism within the population.

Assessment of effect of polymorphisms on PRPF31 transcription

The fragments containing the three polymorphisms of interest were identified and amplified using genomic DNA from individual RP15011 (Fig. 3A). These fragments were tested for
promoter activity to assess the functionality of the promoter region SNPs (Fig. 3B and C).

The presence of differing SNP alleles in fragments e1 and e4 had no effect on luciferase reporter activity ($P > 0.2$). The e3 fragment showed a borderline significant difference between the two individuals in RPE-1 cells ($P = 0.04$), with III.7 showing a slightly higher reporter activity; this effect was not, however, observed in HeLa cells ($P > 0.05$) and suggests that a tissue-specific TF might underlie the slight differences observed in the promoter activity of this fragment.

The absence of the 38 bp duplication and concomitant C allele at rs78558230 did, however, affect the luciferase reporter activity in e7. In fragment e7, these alleles significantly increased reporter activity [2.3-fold increase (RPE-1), 4.5-fold increase (HeLa); $P < 2 \times 10^{-7}$ for both cell lines]. The asymptomatic individual, therefore, harboured the ‘lower-expressor’ haplotype, while the symptomatic individual harboured the ‘higher-expressor’ haplotype. These functional polymorphisms were not, then, investigated further, as the functional effect did not correlate with the clinical phenotype.

**Identification of a patient harbouring a pathogenic PRPF31 5’UTR mutation**

A heterozygous single base pair deletion at the PRPF31 TSS (54618847delC, hg19 co-ordinates) was identified in an adRP patient, CAN493 (Fig. 4A). No other point mutations were found in the PRPF31 gene, and multiplex ligation-dependent probe analysis (MLPA) confirmed that there were no large deletions or insertions in this patient (data not shown). A screen of major adRP genes ($RHO$, $RP1$, $PRPF3$, $IMPDH$ and $TOPORS$) also found no changes.

The initial stages of a systematic procedure for the assessment of changes in a promoter region were carried out (28). The single base pair deletion was not observed in an ECACC control panel of 96 normal individuals and it was
also not reported in the 1000 genome database, strongly sug-
gesting that this change is not a polymorphism. Evolutionary
conservation of the change was assessed by comparison of
orthologous sequence from nine mammalian species, dem-
strating that the base was conserved in primates, including
lower primates, and also dogs and pigs (Fig. 4B). The base

Figure 3. (A) Schematic representation of the fragments assayed to assess functionality of SNPs. The relative position of the three polymorphisms that differed between an asymptomatic and a symptomatic individual are indicated. (B and C) Results of the dual reporter luciferase assay in HeLa (B) and RPE-1 (C) cells, showing the reporter activity in III.7 and RP15011; error bars show 1 standard deviation. Borderline significant differences ($P = 0.04–0.05$) are marked with one asterisk (*), significant differences ($P < 0.04$) are marked with two asterisks (**).
was not, however, conserved in rabbits or mice, although analysis of the phylogenetic relationship of the region in these species (rabbits and mice when compared with other mammals) suggested that there is little similarity across the entire PRPF31 5' region (Fig. 4C).

Clinical history of patient CAN493

CAN493 is a 66-year-old male from a family of French-Canadian origin affected by adRP (Fig. 5A). Segregation analysis was performed in three available siblings of CAN493, showing that the other affected sibling (III.5) carried the same change. Individual III.6 also harboured the change, although he was not clinically affected by RP, individual III.2 did not carry the change; no other siblings were available to test. Microsatellite markers were analysed in CAN493 and III.6 (the asymptomatic mutation-carrying sibling), showing that the two siblings inherited different WT chromosome 19q13 alleles, consistent with the prevailing theory that inheritance of different chromosome 19 WT alleles determines the clinical outcome of PRPF31 mutations in symptomatic–asymptomatic sibling pairs (22).

The patient (CAN493) first presented with nyctalopia and colour vision deficiency, with difficulty with blue-green discrimination. Past medical history was unremarkable, and additional past ophthalmic history included bilateral cataract extraction with posterior chamber intraocular lens implantation. At last examination, aged 65 years, the patient retained good central vision (Snellen visual acuity 20/20 each eye). The intraocular pressure was 15 mmHg in each eye and slit lamp examination of the cornea, anterior chamber and iris was unremarkable. Colour vision testing using Ishihara plates scored 1/17 in the right eye and 1.5/17 in the left eye. The fundi were examined showing classical changes of RP (Fig. 5B–E); Goldmann visual field testing revealed bilateral large temporal and inferior visual field loss that can be correlated to the retinal changes observed (Supplementary Material, Fig. S4). Optical coherence tomography (OCT) and autofluorescence imaging showed intact foveal photoreceptor structures on OCT and relatively mild fundus autofluorescence (FAF) changes, with normal autofluorescence except for a mild perifoveal annular ring of hyperfluorescence and peripheral mottling (Fig. 5F–I).

In vitro functional studies to assess the effect of mutation

Constructs of both the WT and mutant (delC) alleles of the core promoter fragments (Bi-P+ and P.31-Luc−) were produced from the genomic DNA of CAN493 and tested by the dual-luciferase reporter assay in HeLa and RPE-1 cell lines (Fig. 6A and B). WT fragments Bi-P+ and P.31-Luc− showed near-identical activity to the control individual (III.7), with no significant differences observed; this was to
be expected, as the base sequences of the control individual (III.7) and CAN493 WT fragment were identical.

The mutant Bi-P+ fragment (the core promoter of PRPF31), which differed from the control construct only at the deleted C, showed a 56% reduction in promoter activity when compared with the control and WT fragments in both cell lines [3.63 ± 0.65 (HeLa); 3.10 ± 0.49 (RPE-1)]. In the HeLa cell line, median-fold inductions over pTK for the control and mutant were 8.5 and 3.6, respectively; the distributions in the two groups differed significantly (two-tailed Mann–Whitney $U = 110$, $n_1 = 11$, $n_2 = 10$, $P < 1 \times 10^{-5}$).

A similar situation was observed in the RPE-1 cell line, where median-fold inductions for the control and mutant were 6.8 and 3.2, respectively; the distributions showed highly significant differences between the two groups (two-tailed Mann–Whitney $U = 240$, $n_1 = 24$, $n_2 = 10$, $P < 1 \times 10^{-7}$).

The mutant P.31-Luc plasmid (the TFPT core promoter) also showed a significant reduction in promoter activity when compared with the control and WT fragments in both
cell lines, although this was more striking in HeLa cells [0.37 ± 0.08, 82.9% reduction (HeLa); 0.77 ± 0.16, 41.1% reduction (RPE-1)], both differences were statistically significant (two-tailed Mann–Whitney U-test, $P < 1 \times 10^{-5}$).

DISCUSSION

In this study, we have identified and characterized the core promoters controlling the expression of PRPF31 and TFPT, a bi-directional gene-pair in a head-to-head arrangement with a partially shared exon 1. To identify the core promoters, the dual-luciferase reporter assay was performed in HeLa and RPE-1 cell lines.

The core promoter of TFPT was defined as a fragment (P.31-Luc) spanning −78 to +615 relative to the TFPT TSS, with comparatively weak promoter activity (1.3–2.2-fold induction over pTK minimal promoter). The role of TFPT is still not entirely clear, but there is growing evidence that the protein is involved in p53-independent cellular apoptosis (7–9). The weak promoter activity of TFPT might be because cells were transfected in a healthy state and that the transcription factors necessary for TFPT activation are only present under apoptotic conditions. It should also be noted that these results do not agree with the previously reported TFPT promoter, where the core promoter was defined as Fb-Luc (23), this construct having no activity in our assay (0.07–0.31 times the pTK minimal promoter). The previous study (23) did not, however, compare luciferase activity to pTK, only to the pGL2-basic vector (equivalent to our pGL3-basic vector). If a comparison to pGL3-basic was performed for our assay, then the Fb-Luc fragment would show an almost 50-fold induction in HeLa cells, a figure comparable to that reported by Brambillasca et al. (23).

The core promoter of PRPF31 was defined as the fragment Bi-P, spanning −397 to +539 of the annotated PRPF31 TSS, this construct showing 7.0–8.3-fold induction of activity over a minimal pTK promoter. The definition of a strong promoter (active under normal conditions) is not surprising, given that PRPF31 is a ubiquitous, house-keeping gene. The promoter is, however, unusual in its relation to exon 1 of the gene: most of the core promoter is comprised of exon 1 and intron 1, with a relatively small contribution from the 5′-flanking sequence. EMSA demonstrated that the Bi-P fragment was bound in a specific manner by nuclear proteins; further work could be carried out to define which nuclear TF binds the promoter, thereby more fully characterizing the regulatory network of PRPF31.

Figure 6. Results of dual-luciferase reporter assay using CAN493 WT and CAN493 delC plasmids in (A) HeLa and (B) RPE-1 cell lines. The pTK minimal promoter control is shown for comparison; error bars show 1 standard deviation.
Clinical manifestation of adRP in PRPF31 mutation-carrying individuals is determined by differential expression levels of the PRPF31 gene. Studies in many diverse species—including yeast, maize, flies, mice and primates—have shown that there is extensive variation in gene expression levels within natural populations (29–33). The importance of cis-regulatory element SNPs in the evolution of gene expression variation is not well understood, as most knowledge of genome evolution comes from the study of protein-coding sequences. The study of SNPs in coding regions has provided a reliable and powerful method for studying genome evolution, through analytical models such as usage bias, the Ka:Ks ratio and synonymous:non-synonymous substitution ratio (34). There is not, however, a comparable analytic framework for cis-regulatory elements: this is most likely due to the less-ordered arrangement of cis-regulatory modules, with short stretches of functional nucleotides interspersed within non-functional sequences. Additionally, TF–DNA interactions are degenerate, with one TF interacting with several motifs and vice versa; furthermore, these interactions are strongly context-dependent, with cell type- and cell cycle-specific interactions. Unfortunately, there are no robust and reliable computational methods for distinguishing functional regulatory SNPs in non-coding DNA using nucleotide sequence alone (35). 

Comparison of the 5′UTR and upstream region of PRPF31 in a hemizygous asymptomatic and a hemizygous symptomatic individual led to the identification of three SNPs that might play a role in phenotypic non-penetrance, two single base pair changes and a 38 bp insertion. The presence of the 38 bp insertion, with a concomitant minor allele of rs78558230 (‘low-expressor haplotype’), did affect transcriptional activity (reducing luciferase activity 2.3–4.5 times), indicating that these are functional polymorphisms. Due to the close proximity of the two polymorphisms, it was not possible to assess the individual contribution to the change in reporter activity.

Differential expression of PRPF31 is known to underlie phenotypic non-penetrance associated with adRP. We considered that the observed functional polymorphisms might underlie the differential gene expression level. The presence of the ‘low-expressor haplotype’ in the asymptomatic individual disproved this, and as such, although the polymorphisms described here are functional, they cannot be held responsible for phenotypic non-penetration.

It is possible that an intronic regulatory element polymorphism, or a long-range cis-regulatory element polymorphism, might underlie the differential gene expression of PRPF31. It also needs to be considered that multiple regulatory polymorphisms within the 19q13 region might act co-operatively to define the final PRPF31 expression level, and until these multiple factors are defined, it will be impossible to assess the relative risk contributed by each allele.

The characterization of a patient (CAN493) harbouring a single base pair deletion in the core promoter of TFPT and PRPF31 represents the first description of functional haploinsufficiency as a disease mechanism in adRP. It was shown that the patient harboured no pathogenic changes in the open reading frame of PRPF31, but that a single base pair deletion in the core promoter was likely to have a significant adverse effect on transcription of the gene. The mutant core promoter fragment for PRPF31 had luciferase reporter activity of less than half the WT allele, suggesting that in vivo PRPF31 transcription from the mutant chromosome would be significantly reduced, leading to haploinsufficiency. It was also shown that the transcription of TFPT was significantly reduced, although from previous reports of large deletions in the 19q13 region, it is known that haploinsufficiency of TFPT has no phenotypic effect (27,36). The patient had one sibling who harboured the mutation, but was not affected by RP. PRPF31 expression levels were not measured, but we would expect that the PRPF31 mRNA level would be significantly higher in the asymptomatic sibling, compared with the symptomatic sibling, as has been consistently demonstrated in previous studies (22,37,38).

It is predicted that 1% of single base pair changes causing human genetic disease are located within promoter regions and cause disease through the disruption of normal expression of the gene (39). It is possible, however, that this figure represents an underestimate of the importance of regulatory mutations as a cause of human disease, as the regulatory regions of genes are not routinely screened in genetic diagnostic facilities. It is also often difficult to ascertain whether a promoter change is functionally relevant, genetically linked to another causative location, or a mere incidental finding, and so it remains essential to perform functional studies to show an in vitro effect of regulatory region mutations (40). The Canadian family reported here represents the first case of a promoter region mutation in a retinal degeneration, this mutation leading to functional haploinsufficiency. There might, however, be more such cases and, as such, screening of the promoter region of genes implicated in RP would be worthwhile.

In summary, the regulation of gene expression is a complex process, controlled by a multitude of both cis- and trans-acting factors. The definition of the core promoter of a gene represents the first step in understanding the regulation of expression of the gene in health. Normal variations (polymorphisms) within gene promoters can have significant effects on gene expression level, which can underlie variable gene expression levels within the population. It was found, however, that functional promoter polymorphisms were not responsible for the observed variable expression level of PRPF31, and further work will have to be conducted to find the causal factor, or factors. We have shown that disruption of the core promoter element can have serious consequences for human disease, with a single base pair deletion within the PRPF31 promoter causing incurable visual handicap. This case underlines the importance of control of gene expression in both health and disease and emphasizes the need for continued research in this complicated field.

MATERIALS AND METHODS

Fragment design and amplification

Primers were designed to amplify the fragments described by Brambillasca et al. (23) (‘Group 1’), and in addition, nine further overlapping fragments extending into TFPT intron 3 were defined and primers designed (‘Group 2’).
PCR was performed using KOD polymerase (Novagen) under standard conditions (Supplementary Material, Table S1). The template genomic DNA was extracted from peripheral blood lymphocytes of an asymptomatic individual (III.7) known to harbour a 112 kb deletion, encompassing the entire region of interest (27). PCR product was purified using Wizard SV gel and PCR purification kit (Promega, UK).

Cloning into pGL3 vector

pGL3-basic vector (Promega) was digested using SmaI restriction endonuclease (Promega) at 25 °C for 2 h, followed by inactivation at 70 °C for 20 min. Purified PCR product (100 ng) was combined with 50 ng digested pGL3-basic vector, 1 μl T4 DNA ligase (Promega), 1 μl T4 DNA ligase buffer (Promega) and distilled water to a final volume of 10 μl. This mixture was incubated at room temperature for 4–16 h to allow ligation; this being transformed into DH5α competent cells (Invitrogen) and plated on agar with ampicillin. Single colonies were selected and colony PCR performed using vector-specific primers flanking the restriction insertion site. Gel electrophoresis allowed the identification of colonies containing an insert; the colony PCR product from these was purified using ExoSAP-IT (USB), prior to sequencing with BigDye v3.1 (Applied Biosystems) under standard conditions. The insert sequence was verified against a reference genome, and any unreported changes were verified in a control panel of 96 DNAs (ECACC control panel). Those colonies containing correct fragments were propagated in the lysogeny broth medium and the plasmid isolated by maxi-preparation (PureYield™ Plasmid Maxiprep System, Promega). The Group 1 fragments were obtained in both orientations, forward and reverse. The Group 2 fragments were first obtained in forward strand orientation, and if activity was seen on the luciferase assay, the reverse strand orientation fragment was sought.

Dual-luciferase reporter assay

HeLa and RPE-1 cells (ATCC) were grown in 24-well tissue culture plates under normal conditions, until 80–90% confluent. Co-transfection with pGL3-insert (400 ng) and pRenilla (50 ng, Promega) was performed using Lipofectamine-2000 transfection reagent (Invitrogen), each transfection being performed in quadruplicate. The cells were washed after 4 h, and left to grow at 37°C for 48 h. Cell lysis was performed using Passive Lysis Buffer (Promega) and active cell scraping; 20 μl of the cell lysate was transferred to a well of an opaque white 96-well plate (Costar). The dual-luciferase reporter assay was performed as instructed by the kit manufacturer (Promega). For each fragment, at least three separate transfections were performed. On each plate, a modified pGL3 plasmid containing thymidine kinase-minimal promoter (pTK) was used as a positive control; additionally, pGL3-basic vector and WT cells were assayed as a negative control. Data were disregarded if the pRenilla values for repeats varied more than ±10%.

The data were firstly analyzed by calculating the ratio between pGL3-insert and pRenilla; this calculation standardized for cell number and transfection efficiency. This ratio was then compared with the average value of the positive control, pTK, giving a proportional activity compared with pTK. Mean values and standard deviations for the proportional activity were calculated and the statistical significance of observed differences was assessed using the two-tailed Mann–Whitney U-test.

Mutation screening of patients

One patient, CAN493, with confirmed clinical diagnosis of RP was selected, having had full ophthalmic examination (including slit-lamp examination, assessment of visual acuity, perimetry, colour vision) and electrodagnostic testing at McGill University Health Center, Montreal, Canada. Genomic DNA was extracted from peripheral blood leucocytes under standard conditions. Coding region mutations were excluded in the major adRP genes (PRPF3, PRPF31, Rho, TOPORS and RP1) by PCR amplification of exons and sequencing (previously described). MLPA was performed using P235 Retinitis Pigmentosa kit according to the protocol provided by the manufacturer (MRC-Holland). This kit screens for large insertions or deletions in the four most commonly affected adRP genes: RHO, PRPF31, IMPDH1 and RP1. MLPA fragments were separated and sized using 500 LIZ™ size standard (ABI) on 3730 DNA Analyser (ABI) and the results analysed using GeneMarker® v1.8 (SoftGenetics, PA, USA).

Genotyping of microsatellite markers

Four microsatellite markers spanning 19q13 were selected, which flanked the PRPF31 gene (D19s921, D19s572, D19s927, D19s210). Oligonucleotide primers were designed to amplify the microsatellite of interest; the forward primer had a 5′FAM label. A reaction was performed using 6.4 μl GoTaq Hot Start Colorless Master Mix (Promega), 6.4 μl demineralized water, 60 μM forward primer, 60 μM reverse primer and 30 ng genomic DNA (final reaction volume 15 μl). The PCR was run on a thermal cycle [initial denaturation (95°C, 10 min); 35 cycles of 95°C (30 s), 55°C (30 s), 72°C (45 s); then final extension (72°C, 10 min)]. The PCR (2 μl) was diluted with 10 μl HiDi formamide (ABI, UK) and 0.2 μl GeneScan 500 LIZ size standard (ABI). The mixture was denatured at 95°C for 5 min and then placed on ice, before analysis on ABI3730 DNA sequencer. Microsatellite data were analysed on GeneMarker v1.8 (SoftGenetics).

Bioinformatic analyses

The assessment of evolutionary conservation of bases was performed by alignment of multiple species sequence using the MultiTF tool on the Mulan interface (41). TF-binding sites were identified using MatInspector tool online (http://www.genomatix.de/). CpG content was assessed and CpG islands were identified using EMBoss CpGPlot (http://www.ebi.ac.uk/Tools/emboss/cpgplot/).

Electrophoretic mobility shift assay

EMSA was performed to verify the binding of putative promoter region DNA sequence by nuclear transcription factors. A 5′-biotin-labelled fragment (Bi-P) was produced by PCR
REFERENCES


SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* Online.

ACKNOWLEDGEMENTS

The authors would like to thank the patients and their families for participation in this study. Gratitude is also expressed to Prof. K. Matter for the use of luminometer facilities and to Quincy Prescott and Beverley Scott for technical assistance. We also thank all members of the Bhattacharya laboratory and Mr G.E. Rose for insightful discussions and suggestions.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by grants from Rosetrees Trust (A.M.R., grant number M68), The Astor Foundation (A.M.R., studentship), Fight for Sight (A.Z.S., grant number N.H.W., grant number BMRC012), Foundation Fighting Blindness (C.F.C., grant number C-GE-0710-0506-MEHIO-04), RP Fighting Blindness (G.A., grant number GR561), Foundation Fighting Blindness Canada (R.K.K.) and Canadian Institutes for Health Research (R.K.K.).


