European Journal of Human Genetics (2011) 19, 1074–1081

www.nature.com/ejhg

ARTICLE

# A dominant mutation in *RPE65* identified by whole-exome sequencing causes retinitis pigmentosa with choroidal involvement

Sara J Bowne[1,6], Marian M Humphries[2,6], Lori S Sullivan[1,6], Paul F Kenna[2,3,6], Lawrence CS Tam[2], Anna S Kiang[2], Matthew Campbell[2], George M Weinstock[4], Daniel C Koboldt[4], Li Ding[4], Robert S Fulton[4], Erica J Sodergren[4], Denis Allman[2], Sophia Millington-Ward[2], Arpad Palfi[2], Alex McKee[2], Susan H Blanton[5], Susan Slifer[5], Ioanna Konidari[5], G Jane Farrar[2], Stephen P Daiger[1] and Peter Humphries*[,2]

**Linkage testing using Affymetrix 6.0 SNP Arrays mapped the disease locus in TCD-G, an Irish family with autosomal dominant retinitis pigmentosa (adRP), to an 8.8 Mb region on 1p31. Of 50 known genes in the region, 11 candidates, including *RPE65* and *PDE4B*, were sequenced using di-deoxy capillary electrophoresis. Simultaneously, a subset of family members was analyzed using Agilent SureSelect All Exome capture, followed by sequencing on an Illumina GAIIx platform. Candidate gene and exome sequencing resulted in the identification of an Asp477Gly mutation in exon 13 of the *RPE65* gene tracking with the disease in TCD-G. All coding exons of genes not sequenced to sufficient depth by next generation sequencing were sequenced by di-deoxy sequencing. No other potential disease-causing variants were found to segregate with disease in TCD-G. The Asp477Gly mutation was not present in Irish controls, but was found in a second Irish family provisionally diagnosed with choroideremia, bringing the combined maximum two-point LOD score to 5.3. Mutations in *RPE65* are a known cause of recessive Leber congenital amaurosis (LCA) and recessive RP, but no dominant mutations have been reported. Protein modeling suggests that the Asp477Gly mutation may destabilize protein folding, and mutant *RPE65* protein migrates marginally faster on SDS-PAGE, compared with wild type. Gene therapy for LCA patients with *RPE65* mutations has shown great promise, raising the possibility of related therapies for dominant-acting mutations in this gene.**
*European Journal of Human Genetics* (2011) **19,** 1074–1081; doi:10.1038/ejhg.2011.86; published online 8 June 2011

**Keywords:** retinitis pigmentosa; choroideremia; *RPE65*; exome capture; next-generation sequencing

## INTRODUCTION

Retinitis pigmentosa (RP; MIM 268000) is an inherited retinal degeneration caused by progressive loss of rod and cone photoreceptors. Patients experience night blindness followed by loss of visual fields, culminating in legal and often complete blindness. Clinical hallmarks of RP include bone-spicule deposits, attenuated blood vessels, optic disc pallor, visual field loss, and abnormal, diminished or unrecordable electroretinograms (ERG).[1,2] RP affects approximately one in 3200, with an estimated total of 1.5 million affected worldwide.[3]

RP can be inherited in an autosomal recessive, dominant or X-linked manner, with rare digenic and mitochondrial forms. RP is also associated with several syndromic disorders such as Bardet-Biedl and Usher syndrome. Mutations in 54 genes have so far been identified in non-syndromic RP.[4] These encode proteins involved in a range of retinal functions including phototransduction, outer-segment disc structure, transcriptional regulation, pre-mRNA splicing, and *de novo* GTP biosynthesis.[5,6] Despite the large number of RP genes, mutations cannot be identified in 30–35% of patients with adRP.[7] This manuscript details the clinical assessment, linkage mapping, exome capture, and next-generation sequencing used to identify an *RPE65* c.1430G>A, p.Asp477Gly mutation in a large Irish adRP family with a distinctive phenotype.

## MATERIALS AND METHODS

### Clinical assessment and DNA extraction

Human research was approved by the Research Committee, Eye and Ear Hospital, Dublin and Committee for Protection of Human Subjects, University of Texas Health Science Center, Houston. Families were assessed clinically at the Research Foundation, Eye and Ear Hospital Dublin. Best-corrected Snellen visual acuity, Goldmann perimetry and dark-adapted thresholds were measured. Ganzfeld ERG was performed to International Society for Clinical Electrophysiology of Vision (ISCEV) standards.[8] Fundus photographs were taken using Topcon TRC50X or TRC50DX cameras (Topcon Ireland, Dublin, Ireland).

DNA was extracted from blood using either the QIAamp DNA Blood kit or Gentra Puregene blood kit (Qiagen, Hilden, Germany). DNA was subjected to whole-genome amplification (WGA) by Qiagen's REPLI-g genome amplification service. Only WGA DNAs with assessment ratings indicating a >99% accuracy rate were used.

[1]Human Genetics Center, The University of Texas Health Science Center, Houston, TX, USA; [2]Ocular Genetics Unit, Department of Genetics, Trinity College Dublin, Dublin, Ireland; [3]Research Foundation, Royal Victoria Eye and Ear Hospital, Dublin, Ireland; [4]Genome Institute at Washington University School of Medicine, St Louis, MO, USA; [5]Dr. John T. MacDonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA
*Correspondence: Dr P Humphries, Professor of Medical Molecular Genetics, Ocular Genetics Unit, Institute of Genetics, Trinity College Dublin, Lincoln Place Gate, Dublin 2, Ireland. Tel: +353 1 8962484; Fax: +353 1 8963848; E-mail: pete.humphries@tcd.ie
[6]These authors contributed equally to this work.
Received 25 January 2011; revised 4 April 2011; accepted 14 April 2011; published online 8 June 2011

A total of 27 DNAs from the TCD-G family were sent to The Hussman Institute for Human Genomics, Center for Genome Technology for genotyping using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA, USA). DNA was quantified by the ND-8000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and DNA quality was evaluated by agarose gel electrophoresis. Qualifying DNA was digested with NspI and StyI restriction enzymes (New England Biolabs, Ipswich, MA, USA), ligated to adapters and amplified using adapter-specific primers. Purified PCR products were fragmented, labeled and loaded on the SNP 6.0 arrays. Hybridization was done overnight in a GeneChip Hybridization Oven (Affymetrix). Arrays were washed and stained with streptavidin phycoerythrin (SAPE) on the Fluidics Station 450 (Affymetrix). Arrays were scanned on a GeneChip Scanner 3000 7G (Affymetrix).

### Linkage analysis

SNP genotype calls and CNV analysis were performed on raw Affymetrix Genome-Wide Human SNP 6.0 array data using Genotyping Console (Affymetrix). PLINK (http://pngu.mgh.harvard.edu/purcell/plink) was used for quality control and assessment.[9] A subset of SNPs was chosen for linkage analysis on the basis of heterozygosity and inter-SNP distance. All chosen SNPs were at least 0.2 cM apart and had an average heterozygosity of 0.5. CEPH Caucasian (CEU) dataset allele frequencies were used for all calculations.[10] Multipoint linkage analysis was performed using Merlin (http://www.sph.umich.edu/csg/abecasis/Merlin/download) and a dominant model with 90% penetrance in heterozygotes and disease allele frequency of 0.0001.[11,12] Fast Link (ftp://fastlink.nih.gov/pub/fastlink) (TCD-G) and Merlin (TCD-H) were used for two-point linkage analysis of the *RPE65* c.1430 genotype with disease, based on a dominant model with 80% penetrance.

### Fine-scale linkage region

STR markers were selected from the ABI Prism Linkage Mapping Set (Applied Biosystems, Foster City, CA, USA) or were GDB markers chosen from the UCSC Genome Browser (http://genome.ucsc.edu). Genomic or WGA DNA was amplified using fluorescently labeled primers (Applied Biosystems), AmpliTaq Gold 360 Master Mix (Applied Biosystems) and standard cycling conditions. PCR products were separated on a 3500 Genetic Analyzer (Applied Biosystems) and analyzed using GeneMapper V4.1 (Applied Biosystems). Haplotype analysis was performed manually.

### Candidate gene sequencing

Each exon and intron/exon junction of 11 candidate genes (*RPE65, PDE4B, SGIP1, DNAJC6, SERBP1, GNG12, WLS, SFRS11, JAK1, PGM1*, and *L1TD1*) was tested for mutations in-house using fluorescent di-deoxy capillary sequencing and a 3500 Genetic Analyzer (Applied Biosystems) as described previously.[7] Sequence analysis was performed using SeqScape (Applied Biosystems) software.

### GAIIx paired-end library construction

Illumina paired-end libraries were made from 1 $\mu$g of DNA according to manufacturer's protocol (Illumina Inc, San Diego, CA, USA) with slight modifications: (i) DNA was fragmented into sizes ranging from 100–500 bp, using a Covaris S2 DNA Sonicator (Covaris Inc., Woburn, MA, USA) (ii) Illumina adapter-ligated DNA was amplified in a single 50 $\mu$l PCR reaction for five cycles, and (iii) solid phase reversible immobilization bead cleanup was used to purify the PCR amplification and select for fragments 300–500 bp in size.

### Exome capture and next-generation sequencing

Sequencing libraries were hybridized with a customized Agilent SureSelect All Exome Kit v2.0 (Agilent Technologies, Santa Clara, CA, USA), according to manufacturer's protocol.[13] The KAPA SYBR FAST qPCR Kit (KAPA Biosystems, Woburn, MA, USA) was used for library quantification. The amount of library required to generate 180 000 clusters on a single lane of the Illumina GAIIx platform (Illumina) (was determined using qPCR results. Three lanes of 2×100 bp paired-end sequence were generated for each patient library using the SBS Sequencing Kit Ver. 3 (Illumina).

### Alignment and variant identification

Illumina reads were mapped to the Ensembl release 45 version of Human NCBI Build 36, using BWA Ver. 0.5.7 (http://bio-bwa.sourceforge.net/) with soft trimming (−q 5).[14] For each sample, individual lane alignments in BAM format were merged using SAMtools r544 (http://samtools.sourceforge.net).[15] Duplicates were marked in the merged BAM files by the MarkDuplicates class of Picard v1.17 (http://picard.sourceforge.net/index.shtml).[16] Reads with mapping quality of zero, or which were marked as duplicates by Picard, were excluded from further analysis.

Putative SNPs and indels were called in the exome data using VarScan 2 (http://varscan.sourceforge.net) and the following thresholds: coverage $\geq 8\times$, phred base quality $\geq 15$, minimum variant allele frequency $\geq 10\%$, FET *P*-value $< 0.05$.[17] False positives were removed from paralogous alignments, local misalignments, sequencing errors, and other factors by filtering SNVs to remove any with strand bias, read position bias, or multiple high-quality mismatches in supporting reads. Predicted indels were filtered to remove small events around homopolymers, which likely were false positives.

### Variant annotation and tiering

Annotation of filtered high-confidence somatic mutations using gene structure and UCSC annotation information was performed and each mutation assigned to one of four tiers as described previously.[17]

### Analysis of potential disease-causing variants

Confirmation and segregation of potential disease-causing variants identified as part of exome sequencing or candidate gene sequencing were done using fluorescent capillary-based di-deoxy sequencing, as described previously.[7] DNAs from Irish controls were tested using BanI (New England Biolabs) restriction digest and agarose gel separation of an *RPE65* PCR fragment, which contained the region harboring the mutation.

### Variant analysis

PolyPhen (Polymorphism Phenotyping, http://www.bork.embl-heidelberg.de/PolyPhen/) and SIFT (Sorting Intolerant from Tolerant: http://blocks.fhcrc.org/sift/SIFT.html) were used to analyze sequence conservation, chemical change, and likelihood of pathogenicity of the *RPE65* Asp477Gly variant.[18–20] The potential affect on splicing was analyzed using Alternate Splice Site Predictor (http://www.es.embnet.org/~mwang/assp.html).[21]

### Computer modeling

AGGRESCAN (http://bioinf.uab.es/aggrescan) was used to predict folding properties of normal and Asp477Gly mutant *RPE65* protein, and PopMuSic-2.0 (http://babylone.ulb.ac.be/PoPV2) was used to predict thermodynamic protein stability changes using the RCBS Protein Data Bank accession number 3FSN (http://www.pdb.org/). Wild-type (WT) and mutant protein structures were analyzed by submitting peptide sequences to Swiss-model (http://swissmodel.expasy.org). Predicted structures were visualized and evaluated using Swiss-PdbViewer, DeepView (http://spdbv.vital-it.ch/).

### cDNA vectors

WT and Asp477Gly mutant *RPE65* cDNA clones with HindIII and XhoI sites at 5′ and 3′ ends, respectively, were constructed from synthetic oligonucleotides and/or PCR products and site-directed mutagenesis (GENEART, Regensburg, Germany) and cloned into pcDNA3.1+ mammalian expression vector (Invitrogen, Carlsbad, CA, USA).

### Cell culture and transfection

HeLa cells (ATCC # CCL-2) were cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum, 1% L-glutamine (2 mM), sodium pyruvate (2 mM), and 1% streptomycin/penicillin, in a 5% $CO_2$ incubator at 37°C. Cultured cells were passaged regularly with trypsin-EDTA (Invitrogen) to maintain exponential growth. Before 24 h of transfection, $5\times10^5$ cells were plated in each well of a 6-well plate. Plasmid of 1 $\mu$g expressing WT or mutant *RPE65* was transfected into Hela cells, using Lipofectamine 2000 as outlined by manufacturer (Invitrogen).

### Western blots

Isolation and extraction of cytosolic proteins was performed as previously described.[22] Membrane proteins were extracted from cultured cells using ProteoExtract Native Membrane Protein Extraction Kit (Calbiochem, San Diego, CA, USA). Protein concentrations were determined by BCA Protein

Assay Kit (Pierce, Rockford, IL, USA). Total protein (30 $\mu$g) was separated by electrophoresis on 10% SDS-polyacrylamide gels run under reducing conditions, and transferred to nitrocellulose membranes.

Membranes were blocked for 1 h, incubated overnight with mouse *RPE65* monoclonal primary antibody (Abcam, Cambridge, UK; 1:5000), washed in TBS, and incubated with horse radish peroxidase-conjugated goat polyclonal mouse IgG secondary antibody (Abcam; 1:2500). Membranes were developed using an Enhanced Chemiluminescent kit (Pierce) and exposed to X-ray film. Membranes were stripped with Restore Western Blot Stripping Buffer (Pierce) and probed with rabbit polyclonal antibody to $\beta$-actin.

## RESULTS

### TCD-G clinical assessment

TCD-G is an Irish family with adRP and signs of choroidal involvement (Figure 1). Although age of onset varies from 2nd to 5th decade, initial presenting symptoms are usually impaired dark adaptation. One individual (Figure 1a, #47), has an affected parent and son but did not show symptoms at age 52.

As disease progresses affected individuals note impairment of peripheral visual fields. Typically, this affects the mid-peripheral field, as shown in Goldmann perimetry records of three individuals (Figures 2a–c). In one mildly affected individual with a paucity of retinal changes, but significantly impaired ERG responses, mid-peripheral field loss was ascribed to low-tension glaucoma, resulting in drainage surgery. Although some individuals develop posterior subcapsular lens opacities in later life, the occurrence of these changes is less common in TCD-G compared with other adRP families.

Funduscopic pictures seen in affected individuals vary widely, depending on severity of symptoms. In mildly affected individuals, the retina appears superficially normal, with normal-appearing optic disc, macula, and retinal vasculature, but with intra-retinal pigmentary deposits in the mid-periphery in all four quadrants. These deposits are mainly of the bone-spicule variety, typical of RP, but some are more 'clumpy' and better described as 'nummular' (Figure 2d). More severely affected individuals show a more dramatic fundus picture, typified by extensive diffuse chorioretinal atrophy (Figures 2e and f). In contrast to widespread chorioretinal atrophy, pigmentary deposits are relatively sparse and typically of the nummular variety.

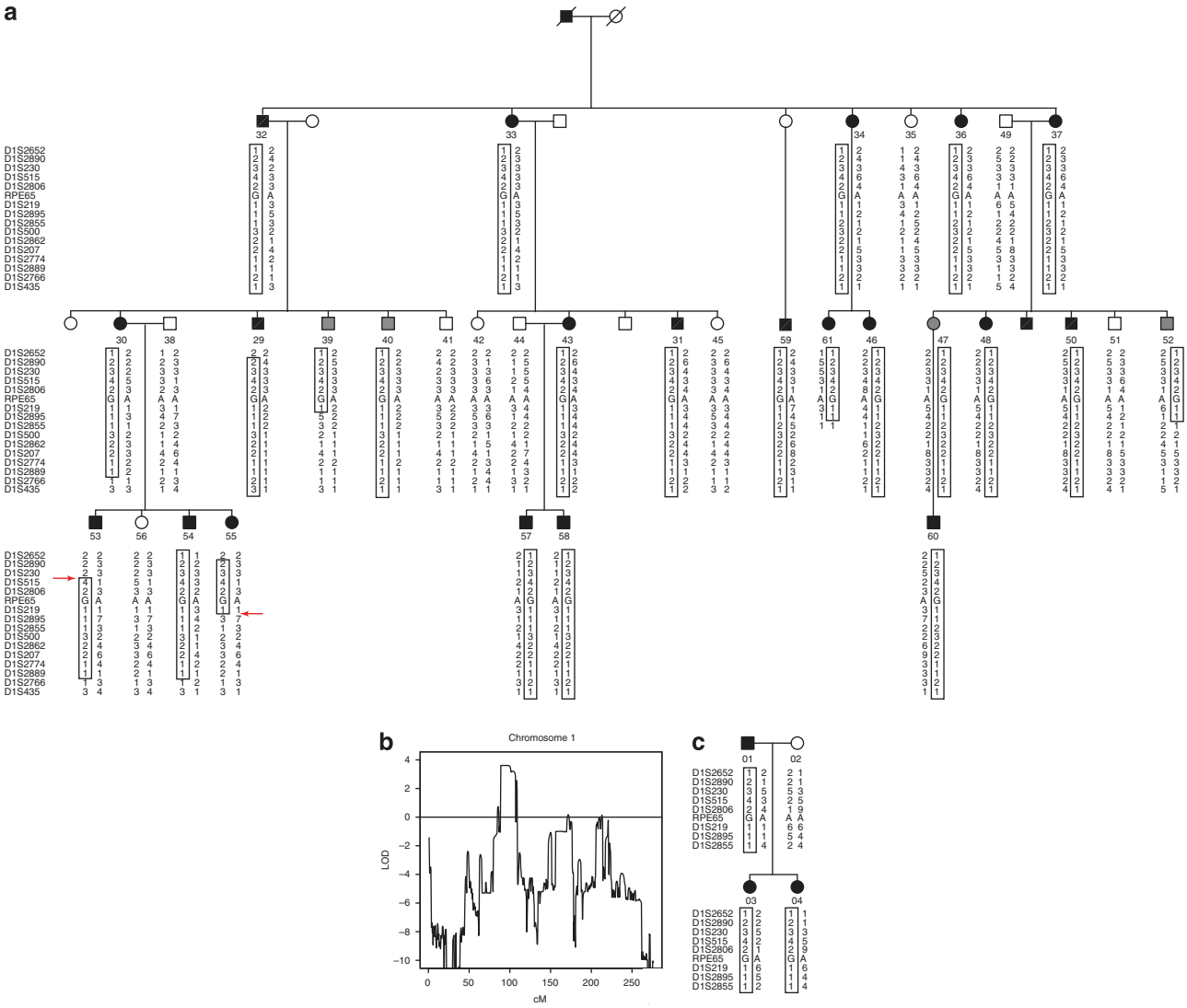


**Figure 1** Linkage analysis in TCD-G (**a**) TCD-G pedigree and STR typing data for linkage region on chromosome 1. (**b**) TCD-G genome wide linkage analysis results for chromosome 1. (**c**) TCD-H pedigree and STR typing data.

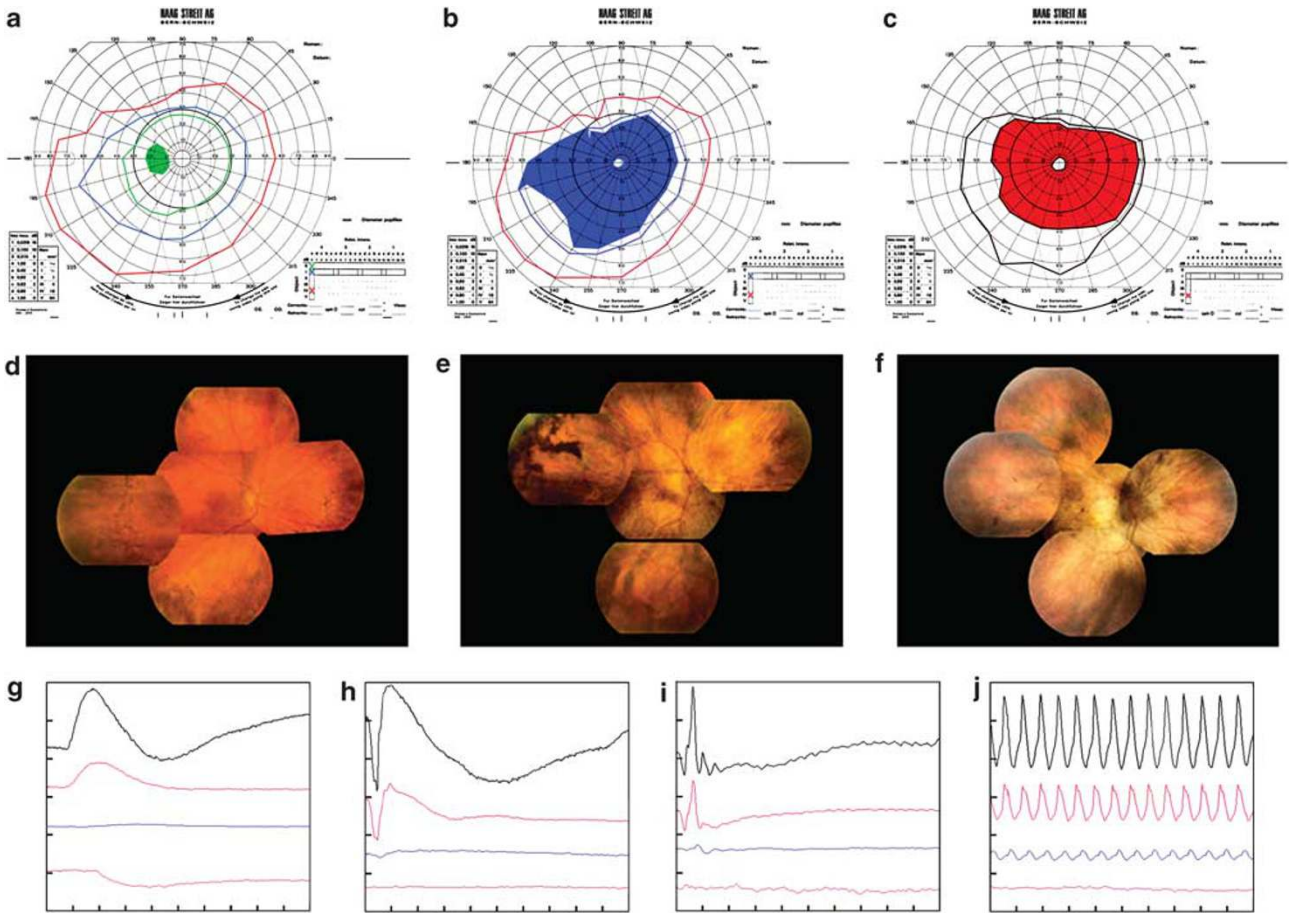


**Figure 2** TCD-G clinical evaluation (**a**) Goldmann perimetry for the right eye from a very mildly affected individual, showing slight loss of superior field to the large IV4e target, (**b**) a moderately severely affected individual illustrating the complete mid-peripheral ring scotoma to the intermediate sized I4e target, and (**c**) a severely affected individual illustrating a complete mid-peripheral ring scotoma to the large IV4e target. (**d**) Composite fundus photographs from the right eye of a mildly affected individual showing bone spicule and nummular intraretinal pigmentary deposits in the mid-periphery. The area of pigment deposition also shows thinning of the underlying retinal pigment epithelium, (**e**) a moderately severely affected individual showing extensive areas of chorioretinal atrophy with pigment clumping particularly evident temporally, and (**f**) a severely affected individual. Extensive, diffuse chorioretinal atrophy is evident, affecting in particular the macular area. Nummular intraretinal pigmentary deposits are visible. This fundus picture bears a superficial resemblance to that seen in choroideremia. (**g**) Rod-isolated electroretinographic responses from a normal individual (black), a mildly affected patient from TCD-G (red), a moderately severely affected patient (blue), and a severely affected patient (purple). Note the reduction in amplitude, but normal timing of the response, from the mildly affected patient compared with the markedly reduced amplitude, and delayed response from the moderately affected person, and the non-recordable response from the severely affected individual. The late-downward deflection in the purple tracing is an artefact due to the patient blinking. The vertical amplitude scale is 250 $\mu$V per division. The horizontal timescale is 50 ms per division. (**h**) Mixed rod and cone electroretinographic responses from a normal individual (black), a mildly affected patient from TCD-G (red), a moderately severely affected patient (blue), and a severely affected patient (purple). Note the reduction in amplitude, but normal timing of the response, from the mildly affected patient compared with the markedly reduced amplitude, and delayed response from the moderately affected person, and the non-recordable response from the severely affected individual. The red and blue responses show relatively better preservation of the negative a-wave compared with the later positive b-wave. The vertical amplitude scale is 250 $\mu$V per division. The horizontal timescale is 50 ms per division. (**i**) 0.5 Hz cone electroretinographic responses from a normal individual (black), a mildly affected patient from TCD-G (red), a moderately affected patient (blue), and a severely affected patient (purple). Note the reduction in amplitude, but normal timing of the response, from the mildly affected patient compared with the markedly reduced amplitude, and delayed response from the moderately affected person, and the non-recordable response from the severely affected individual. The vertical amplitude scale is 100 $\mu$V per division. The horizontal timescale is 50 ms per division. (**j**) 30 Hz cone electroretinographic responses from a normal individual (black), a mildly affected patient from TCD-G (red), a moderately severely affected patient (blue), and a severely affected patient (purple). Note the reduction in amplitude, but normal timing of the response, from the mildly affected patient compared with the markedly reduced amplitude, and delayed response from the moderately affected person, and the non-recordable response from the severely affected individual. The vertical amplitude scale is 100 $\mu$V per division. The horizontal timescale is 50 ms per division.

ERG analysis in affected members of the extended TCD-G family showed impairment of rod function, even in mildly affected individuals. In more severely affected individuals, rod-isolated responses are either severely attenuated or non-recordable (Figure 2g). A similar pattern is evident in mixed rod and cone response recorded from the dark-adapted eye to the maximal intensity flash. This response, in some individuals, had an almost 'electronegative' conformation, with greater diminution of the b-wave component compared with a-wave (Figure 2h). Cone dominated responses to 0.5 and 30 Hz flicker are illustrated in Figure 2(i and j).

### Mutation analysis

No disease-causing mutations were identified when DNA from the TCD-G proband was sequenced for mutations in known adRP genes.[23–27] MLPA analysis of the *PRPF31* gene also failed to identify copy number variants.[28] Recently, two affected members of TCD-G were tested as part of a 1000 amplimer, 46 gene pilot next-generation sequencing project.[7] No disease-causing mutations were identified in TCD-G (also called VCH026).

### Linkage analysis

Using Affymetrix 6.0 genotypes (Affymetrix), multipoint linkage analysis was conducted on TCD-G. This analysis identified a region on chromosome 1p31, linked to disease with a maximum multipoint LOD score of 3.6 (Figure 1b). This region is flanked by rs2182241 and rs7522851, and spans approximately 20 Mb. No significant LOD scores were obtained for any other regions of the genome.

Haplotype analysis of 20 STR markers located on chromosome 1p31 was performed to confirm and refine the critical disease region (Figure 1a). Two recombination events, between D1S230 and D1S515 in TCD-G-53 and between D1S219 and D1S2895 in TCD-G-55, were observed. These recombination events, combined with SNP haplotype analysis, reduced the disease-interval to an 8.8 Mb region located between rs3861941 and D1S2895. The disease haplotype located between these markers is the same for all 20 affected family members tested, and was also observed in three unaffected family members. The presence of affected haplotypes in three individuals without RP symptoms (TCD-G-39, TCD-G-40, and TCDG-52) is consistent with the wide range of clinical phenotypes observed in this family and two documented instances of affected offspring of unaffected, at-risk family members (TCD-G-59 and TCD-G-60).

### Candidate genes

According to UCSC hg18/March2006 build (http://genome.ucsc.edu/cgi-bin/hgGateway), 50 genes are located in the 8.8 Mb critical disease interval on 1q31. Genes were targeted as likely candidates on the basis of: (i) association with other types of inherited retinal degeneration, (ii) similarity to genes associated with inherited retinal degeneration (iii) inclusion in the sensory cilium proteome or EyeSAGE data, or (iv) high expression levels in retina and/or RPE.[4,28,29] Each exon and flanking intron/exon junction of the candidate genes selected (*RPE65, PDE4B, SGIP1, DNAJC6, SERBP1, GNG12, WLS, SFRS11, JAK1, PGM1,* and *L1TD1*) were tested in two affected family members, using fluorescent di-deoxy capillary sequencing. Potential disease-causing variants found in both affected individuals sequenced were tested in the remainder of TCD-G family. The only potential disease-causing variant tracking with disease was an *RPE65* c.1430G>A, p.Asp477Gly (NM_000329) in exon 13.

### Exome sequencing

Concurrent with candidate gene sequencing, three affected and one unaffected TCD-G family members underwent Agilent SureSelect exome capture and next generation sequencing on a GAIIx platform. For each sample, approximately 18 Gbp of sequence data was generated, 98.5% of which was mapped to hs36 with a 13.4% duplication rate (Table 1). On average, 81.5% of the 33 Mbp target coding sequence (CDS) was covered at >20× depth, with 87.1% being covered at >10× and 96.5% being covered at >1×.

Variant calling identified thousands of variants (Table 2), separated into four tiers based on location and relevance to alteration of the proteome. Tier 1 variants were most likely to alter CDSs of genes and hence were the subject of further analysis. This tier included exonic nonsynonymous and synonymous variants, splice-site variants, and non-coding RNA variants.

A total of 3437 tier 1 variants were found in one or more of the three affected samples sequenced. These variants were further prioritized by removing any identified in the unaffected (control) DNA sample (1507) or variants believed to be synonymous (557), using the assumption that protein-altering variants are more likely to be the cause of disease. The remaining 1373 genomic variants were then restricted to those located in the linkage region on chromosome 1. The three remaining variants located in the critical disease region were an A>G change in noncoding RNA LOC645291, a C>T change in the intronic splice region of *ANKRD13C*, and a T>C change in *RPE65* predicted to cause an amino acid substitution (Asp477Gly). Of these three variants, the only one identified in all three affected individuals was the *RPE65* c.1430G>A, p.Asp477Gly, which was the same variant as that identified during candidate gene screening.

Mutations in *RPE65* are a known cause of autosomal recessive RP and LCA, but have not previously been associated with dominant disease. Exome capture and sequencing analyzed 97% of the CDS found in the linkage region at a depth >20×. To minimize the possibility that the *RPE65* c.1430 G>A variant was in linkage disequilibrium with a true disease-causing variant, the remaining 3% of CDS regions located in the critical disease interval were analyzed using di-deoxy sequencing. No potential disease-causing variants were identified.

**Table 1 Exome data summary**

| | |
|---|---|
| Number of samples | 4 |
| Sequencing platform | Illumina GAIIx |
| Read length | 2×100 bp |
| Lanes per sample | 3 |
| Sequence per sample | 18 Gbp |
| Mapping rate (Hs36) | 98.52% |
| Duplication rate | 13.39% |
| Exome platform | SureSelect v2 |
| Genes targeted | 18 500 |
| Exons targeted | 180 000 |
| CDS bases targeted | 33 Mbp |
| Target bases covered ≥20× | 81.53% |
| Target bases covered ≥10× | 87.08% |
| Target bases covered ≥1× | 96.49% |

**Table 2 Exome SNP analysis**

| *Sample* | *VCH026-01* | *VCH026-02* | *VCH026-03* | *VCH026-04* |
|---|---|---|---|---|
| *status* | *Affected* | *affected* | *unaffected* | *affected* |
| *A. Summary of Mendelian analysis for SNPs* | | | | |
| SNPs passing filters | 93 575 | 100 450 | 97 462 | 99 620 |
| Novel to dbSNP | 9743 | 11 153 | 10 257 | 10 858 |
| Annotated as tier1 | 1655 | 1671 | 1819 | 1764 |
| *B. Unique SNPs in 1+ affecteds* | | | | |
| Annotated tier 1 | 3437 | | | |
| Not called in control | 1930 | | | |
| Non-silent | 1373 | | | |
| Map to linkage region | 3 | | | |

### RPE65 analysis in controls and other retinal degenerations

Ethnically-matched control DNAs were examined for the presence of the *RPE65* Asp477Gly variant. The variant was not detected in 684 Irish chromosomes, further strengthening the likelihood that the Asp477Gly mutation is the cause of RP in TCD-G.

DNAs from 12 Irish patients with a range of inherited retinal degenerations were also tested for the presence of the *RPE65* Asp477Gly mutation. One male patient (TCD-H), initially diagnosed with choroideremia, but without a mutation in the *CHM* gene, was found to have the mutation. Subsequent testing showed that the proband's two affected daughters carried this same mutation. Typing and comparison of STR genotypes surrounding the Asp477Gly mutation on chromosome 1p31 showed that the mutation occurred on the same haplotype in TCD-G and the TCD-H choroideremia family (Figure 1c). Two-point linkage analysis of the *RPE65* Asp477Gly variant with disease in both families gives a combined two-point LOD score of 5.3 at 0% recombination. Sequence analysis of DNA from a previously described adRP cohort[23–25,27] failed to find the Asp477Gly mutation or any other pathogenic variant in *RPE65*. This suggests that dominant RPE mutations are a rare cause of retinal degeneration with choroidal involvement.

**Table 3 Conservation of amino acid change (Asp477Gly)**

| *Organism* | *Amino acid sequence* |
|---|---|
| Human | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Chimp | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Macaque | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Rhesus | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Pig | LCKLNVKTKETWVWQEPDSYPSEPIFVAHPDALEEDD |
| Mouse | LCKMNVKTKEIWMWQEPDSYPSEPIFVSQPDALEEDD |
| Rat | LCKLNVKTKEIWMWQEPDSYPSEPIFVSQPDALEEDD |
| Cow | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Cat | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Dog | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Elephant | LCKLNVKTKETWVWQEPDSYPSEPIFVSHPDAMEEDD |
| Chicken | –CKLNVKTKETWVWQEPDSYPSEPIFVSHPDALEEDD |
| Zebrafish | FCKLNVKSKETWIWQEPDAYPSEPLFVQSPDAEDEDD |

### Protein analyses

The Asp477Gly variant was analyzed *in silico* to determine the likelihood of it being a pathogenic mutation. With a Grantham score of 94, the amino acid change could possibly be significant, though PolyPhen2 and SIFT predict the amino acid change to be benign. The amino acid change occurs at a relatively well-conserved position in the *RPE65* protein, and the aspartic acid is found in all mammals and zebrafish (Table 3). At the nucleotide level, splice site predictor analyses suggest that the change from A to G increases the strength of an alternate splice acceptor site in the middle of the exon, and thus could change the pattern of splicing.

To further examine the possibility that the Asp477Gly mutation might affect folding of the *RPE65* protein, both WT and mutant peptide sequences were submitted to AGGRESCAN. This indicated that WT *RPE65* peptide possesses 23 potential 'hotspot areas' (denoting aggregation-prone segments), whereas an additional hotspot was observed in mutant *RPE65* peptide (Figure 3a). The

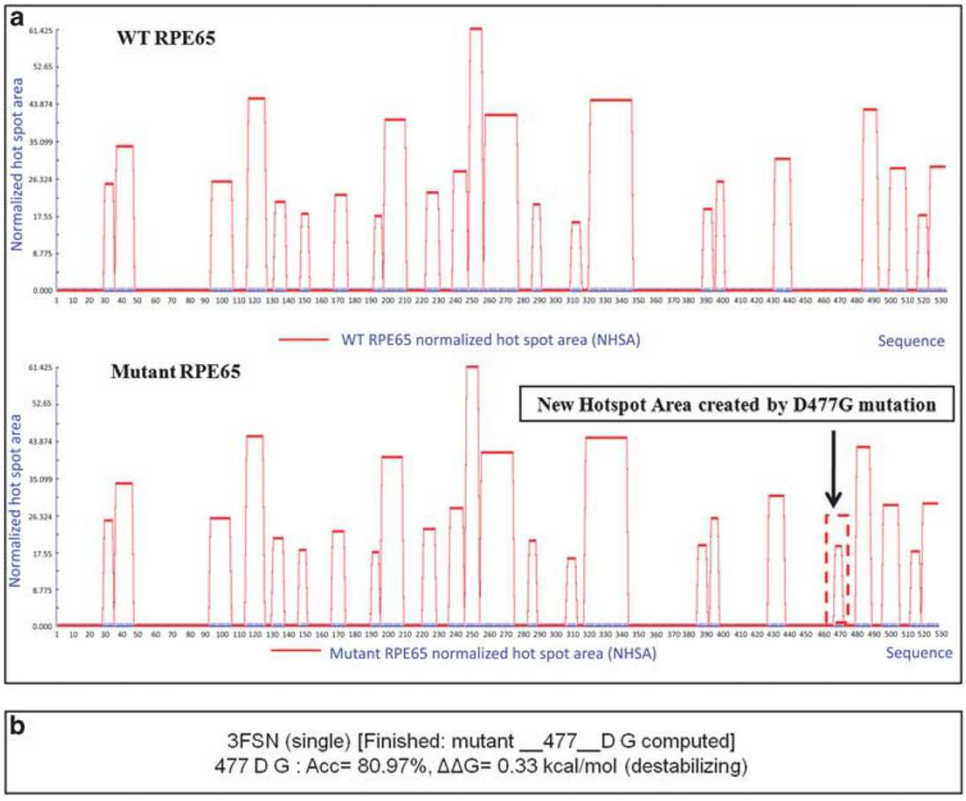


**Figure 3** Computer modeling of mutant *RPE65* protein. (**a**) Hotspot area graphics predicted by AGGRESCAN. Hotspot area plots for WT and mutant *RPE65*, bearing the Asp477Gly mutation. Red peaks represent sequence stretches with highest predicted aggregation propensity. New hotspot area (five amino acids–IFVSH) created by Asp477Gly mutation along the *RPE65* polypeptide sequence. (**b**) PopMuSic V2.0 prediction of protein mutant stability changes. An amino acid change from aspartic acid to glycine at position 477 is predicted to impose a destablizing effect on *RPE65* protein.

additional hotspot contains a stretch of five amino acids (IFVSH), situated in blade VI directly upstream of the Asp477Gly missense mutation. This suggests that this particular missense mutation may have a propensity to perturb normal protein folding.

PopMuSic-2.0 was also used to predict thermodynamic protein stability changes caused by the Asp477Gly mutation. This algorithm indicates that changing amino acid 477 from aspartic acid to glycine yields high-folding free energy ($\Delta\Delta G$=0.33 kcal/mol) that has a destabilizing effect on *RPE65* protein structure (Figure 3b). Furthermore, the mutant tertiary structure generated by Swiss-model predicts that a change from aspartic acid to glycine at amino acid position 477 causes significant perturbations to folding in specific regions of $\beta$-propeller blades VI and VII (Figure 4a), which may have a negative downstream effect on iron coordination. One of the four-His residues required for iron binding, residue 527, is predicted to lie in a distorted region of blade VII of the mutant protein (Figure 4b).

### *In vitro* mutant protein analysis

SDS-PAGE shows that mutant *RPE65* protein migrates marginally faster than WT (Figure 4c). The shift observed for mutant protein may be due to post-translational modification, or to proteolytic cleavages. The palmitoylation of specific cysteine residues in *RPE65* has been shown to be essential for membrane association.[30] Western blot analysis indicates that both WT and mutant *RPE65* expression remain unchanged in membrane fractions (Figure 4d), suggesting that loss of palmitoyl residue is not likely to be the cause of the band shift.

## DISCUSSION

Linkage mapping in an Irish family (TCD-G) with adRP and choroidal involvement mapped the disease locus to an 8.8 Mb region on 1p31. Earlier testing had excluded all known adRP genes as a cause of disease.[7] A combination of candidate gene screening and whole-exome sequencing identified an Asp477Gly mutation in *RPE65* tracking with disease in the family. All 20 affected family members tested have the *RPE65* mutation, as do three unaffected, at-risk family members and one unaffected parent of an affected offspring (an obligate carrier). Sequencing of coding regions and intron–exon junctions of the remaining genes in the linkage region failed to detect any additional, potential disease-causing variants.

The *RPE65* mutation was not detected in 342 Irish controls, but was in a small Irish family (TCD-H), with a diagnosis of choroideremia. The X-linked choroideremia gene, *CHM*[31–33], was excluded as the cause of disease in TCD-H. Three affected family members in TCD-H have the Asp477Gly mutation, and the mode of inheritance and clinical phenotype is consistent with those observed in TCD-G. Furthermore, identical 1p31 haplotypes segregating with disease in both families indicate that the mutation arose in a common ancestor. Combining the two families brings the two-point maximum LOD score to 5.3 at 0% recombination.

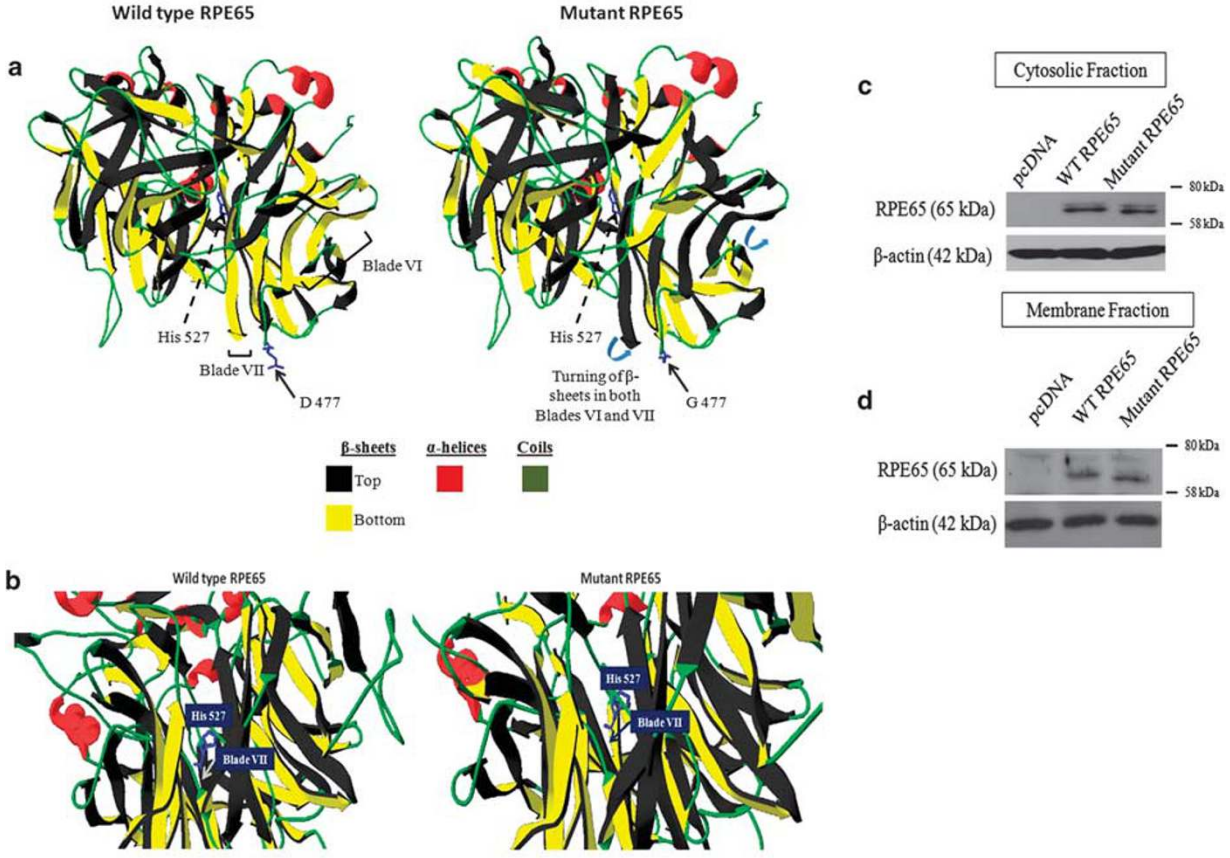


**Figure 4** Predicted and observed properties of mutant *RPE65* protein. (**a**) WT and mutant *RPE65* tertiary protein structures generated by Swiss-model. (**b**) Superimposed images of His 527 residue on WT and mutant *RPE65* secondary protein structures. Blade VII is turned in mutant *RPE65*, following the introduction of Asp477Gly mutation, as compared with WT protein. (**c**) Expression of WT and mutant *RPE65* proteins in the cytosolic fraction: A band shift was detected in mutant *RPE65* protein as compared with WT *RPE65*, which was detected at around 65 kDa. pcDNA as negative control. $\beta$-actin as loading control. (**d**) Expression of WT and mutant *RPE65* proteins in the membrane fraction; both WT and mutant *RPE65* were detected in the membrane fraction, suggesting that the missense mutation did not affect the association of *RPE65* with the cell membrane.

This is the first report of a dominant-acting mutation in *RPE65*; previously reported mutations in *RPE65* are associated with recessive RP or recessive Leber congenital amaurosis.[34,35] The less severe phenotype and occasional instances of non-penetrance in TCD-G are distinct from the severe, completely-penetrant disease seen in recessive *RPE65* cases, and are consistent with one mutant allele rather than two. In spite of genetic and clinical evidence in support of this mutation as the cause of disease in these families, the possibility of a mutation in the linkage region outside of the known genes cannot be completely excluded, though this is unlikely. PolyPhen and Grantham scores for pathogenicity of the Asp477Gly mutant are equivocal; however, the aspartic acid allele at codon 477 is conserved in mammals and zebrafish. In addition, protein modeling suggests that the Asp477Gly mutation may destabilize folding, and expressed mutant *RPE65* protein shows altered mobility from WT. Currently, no clear consensus has emerged regarding *RPE65* palmitoylation, studies having suggested that *RPE65* does not undergo significant post-translational modifications.[36] However, this will need to be analyzed in greater detail before definite conclusions can be reached. Finally, *RPE65* is the most biologically plausible candidate gene in the linkage region.

These data strongly support the conclusion that the *RPE65* Asp477Gly mutation is the cause of disease in these families. This is the first report of a dominant-acting mutation in *RPE65*, and suggests that *RPE65* may harbor other dominant-acting mutations. It also suggests that carriers of 'recessive' missense mutations in *RPE65* should be evaluated for subtle signs of disease. In addition, the frequent co-occurrence of choroidal disease in the large adRP family, and the diagnosis of choroiderema in the smaller family, suggest that mutations in *RPE65* may be the cause of disease in choroideremia families in which the typical, X-linked gene, *CHM*, has been excluded. Finally, AAV-mediated gene-replacement therapy has successfully treated patients with recessive Leber congenital amaurosis caused by *RPE65* mutations.[37–39] Similar approaches may show promise in treating retinal disease caused by dominant-acting mutations in *RPE65*, especially when coupled with targeted suppression and gene-replacement therapies.[40–42]

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We are grateful to TCD-G and TCD-H members who participated in this study, also Mr Hugh Nolan for clinical photography and Ms Hilary Dempsey for electroretinography. This work was supported by The Foundation Fighting Blindness, NIH EY007142, Fighting Blindness-Ireland, Science Foundation Ireland and Health Research Board of Ireland.

1 Heckenlively JR: *Retinitis Pigmentosa*. Philadelphia, PA, USA: J.B. Lippincott, 1988.
2 Heckenlively JR, Daiger SP: *Hereditary Retinal and Choroidal Degenerations*, 5th edn. New York, NY, USA: Churchill Livingston Elsevier, 2007.
3 Haim M: Epidemiology of retinitis pigmentosa in Denmark. *Acta Ophthalmol Scand Suppl* 2002; **233**: 1–34.
4 RetNet: The Retinal Information Network http://www.sph.uth.tmc.edu/RetNet/ Stephen P. Daiger, PhD, Administrator, The Univ. of Texas Health Science Center at Houston, 1996-present.
5 Daiger SP, Bowne SJ, Sullivan LS: Perspective on genes and mutations causing retinitis pigmentosa. *Arch Ophthalmol* 2007; **125**: 151–158.
6 Kennan A, Aherne A, Humphries P: Light in retinitis pigmentosa. *Trends Genet* 2005; **21**: 103–110.
7 Bowne SJ, Sullivan LS, Koboldt DC *et al*: Identification of disease-causing mutations in autosomal dominant retinitis pigmentosa (adRP) using next-generation DNA sequencing. *Invest Ophthalmol Vis Sci* 2011; **52**: 494–503.
8 Marmor MF, Fulton AB, Holder GE, Miyake Y, Brigell M, Bach M: ISCEV Standard for full-field clinical electroretinography (2008 update). *Doc Ophthalmol* 2009; **1**: 69–77.
9 Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
10 HapMap: The International HapMap Project. *Nature* 2003; **426**: 789–796.
11 Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.
12 Cook Jr EH: Merlin: faster linkage analysis with improved genotyping error detection. *Pharmacogenomics J* 2002; **2**: 139–140.
13 Gnirke A, Melnikov A, Maguire J *et al*: Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 2009; **27**: 182–189.
14 Li H, Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; **26**: 589–595.
15 Li H, Handsaker B, Wysoker A *et al*: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
16 McKenna A, Hanna M, Banks E *et al*: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–1303.
17 Koboldt DC, Chen K, Wylie T *et al*: VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 2009; **25**: 2283–2285.
18 Ramensky V, Bork P, Sunyaev S: Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 2002; **30**: 3894–3900.
19 Ng PC, Henikoff S: SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; **31**: 3812–3814.
20 Grantham R: Amino acid difference formula to help explain protein evolution. *Science* 1974; **185**: 862–864.
21 Wang M, Marin A: Characterization and prediction of alternative splice sites. *Gene* 2006; **366**: 219–227.
22 Tam LC, Kiang AS, Campbell M *et al*: Prevention of autosomal dominant retinitis pigmentosa by systemic drug therapy targeting heat shock protein 90 (Hsp90). *Hum Mol Genet* 2010; **19**: 4421–4436.
23 Sullivan LS, Bowne SJ, Birch DG *et al*: Prevalence of disease-causing mutations in families with autosomal dominant retinitis pigmentosa (adRP): a screen of known genes in 200 families. *Invest Ophthalmol Vis Sci* 2006; **47**: 3052–3064.
24 Gire AI, Sullivan LS, Bowne SJ *et al*: The Gly56Arg mutation in NR2E3 accounts for 1–2% of autosomal dominant retinitis pigmentosa. *Mol Vis* 2007; **13**: 1970–1975.
25 Bowne SJ, Sullivan LS, Gire AI *et al*: Mutations in the TOPORS gene cause 1% of autosomal dominant retinitis pigmentosa (adRP). *Mol Vis* 2008; **14**: 922–927.
26 Friedman JS, Ray JW, Waseem N *et al*: Mutations in a BTB-Kelch protein, KLHL7, cause autosomal-dominant retinitis pigmentosa. *Am J Hum Genet* 2009; **84**: 792–800.
27 Sullivan LS, Bowne SJ, Seaman CR *et al*: Genomic rearrangements of the PRPF31 gene account for 2.5% of autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* 2006; **47**: 4579–4588.
28 Liu Q, Tan G, Lenenkova N, Rux JJ, Speicher DW, Pierce EA: The proteome of a mammalian sensory cilium, the mouse photoreceptor outer segment. *Invest Ophthalmol Vis Sci* 2006; E-Abstract: 3725.
29 Bowes Rickman C, Ebright JN, Zavodni ZJ *et al*: Defining the human macula transcriptome and candidate retinal disease genes using EyeSAGE. *Invest Ophthalmol Vis Sci* 2006; **47**: 2305–2316.
30 Takahashi Y, Moiseyev G, Ablonczy Z, Chen Y, Crouch RK, Ma JX: Identification of a novel palmitylation site essential for membrane association and isomerohydrolase activity of RPE65. *J Biol Chem* 2009; **284**: 3211–3218.
31 Cremers FP, van de Pol DJ, van Kerkhoff LP, Wieringa B, Ropers HH: Cloning of a gene that is rearranged in patients with choroideraemia. *Nature* 1990; **347**: 674–677.
32 Cremers FP, Sankila EM, Brunsmann F *et al*: Deletions in patients with classical choroideremia vary in size from 45 kb to several megabases. *Am J Hum Genet* 1990; **47**: 622–628.
33 van Bokhoven H, van den Hurk JA, Bogerd L *et al*: Cloning and characterization of the human choroideremia gene. *Hum Mol Genet* 1994; **3**: 1041–1046.
34 Marlhens F, Bareil C, Griffoin JM *et al*: Mutations in RPE65 cause Leber's congenital amaurosis. *Nat Genet* 1997; **17**: 139–141.
35 Morimura H, Fishman GA, Grover SA, Fulton AB, Berson EL, Dryja TP: Mutations in the RPE65 gene in patients with autosomal recessive retinitis pigmentosa or leber congenital amaurosis. *Proc Natl Acad Sci USA* 1998; **95**: 3088–3093.
36 Yuan Q, Kaylor JJ, Miu A, Bassilian S, Whitelegge JP, Travis GH: Rpe65 isomerase associates with membranes through an electrostatic interaction with acidic phospholipid headgroups. *J Biol Chem* 2010; **285**: 988–999.
37 Bainbridge JW, Smith AJ, Barker SS *et al*: Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* 2008; **358**: 2231–2239.
38 Simonelli F, Ziviello C, Testa F *et al*: Clinical and molecular genetics of Leber's congenital amaurosis: a multicenter study of Italian patients. *Invest Ophthalmol Vis Sci* 2007; **48**: 4284–4290.
39 Hauswirth WW, Aleman TS, Kaushal S *et al*: Treatment of leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. *Hum Gene Ther* 2008; **19**: 979–990.
40 O'Reilly M, Palfi A, Chadderton N *et al*: RNA interference-mediated suppression and replacement of human rhodopsin *in vivo*. *Am J Hum Genet* 2007; **81**: 127–135.
41 Kiang AS, Palfi A, Ader M *et al*: Toward a gene therapy for dominant disease: validation of an RNA interference-based mutation-independent approach. *Mol Ther* 2005; **12**: 555–561.
42 Tam LC, Kiang AS, Kennan A *et al*: Therapeutic benefit derived from RNAi-mediated ablation of IMPDH1 transcripts in a murine model of autosomal dominant retinitis pigmentosa (RP10). *Hum Mol Genet* 2008; **17**: 2084–2100.