Expanding the clinical, allelic, and locus heterogeneity of retinal dystrophies

Nisha Patel, PhD[1], Mohammed A. Aldahmesh, PhD[1], Hisham Alkuraya, MD[2,3], Shamsa Anazi, MSc[1], Hadeel Alsharif, BSc[1], Arif O. Khan, MD[1,4], Asma Sunker, BSc[1], Saleh Al-mohsen, MD[5], Emad B. Abboud, MD[6], Sawsan R. Nowilaty, MD[6], Mohammed Alowain, MD[7], Hamad Al-Zaidan, MD[7], Bandar Al-Saud, MD[5], Ali Alasmari, MD[8], Ghada M.H. Abdel-Salam, MD[9], Mohamed Abouelhoda, PhD[1,10], Firdous M. Abdulwahab, BSc[1], Niema Ibrahim, BSc[1], Ewa Naim, BPharm[1,10], Banan Al-Younes, MSc[1,10], Abeer E. AlMostafa, BSc[1,10], Abdulelah AlIssa, BSc[1,10], Mais Hashem, BSc[1], Olga Buzovetsky, PhD[11], Yong Xiong, PhD[11], Dorota Monies, PhD[1,10], Nada Altassan, PhD[1,10], Ranad Shaheen, PhD[1], Selwa A.F. Al-Hazzaa, MD[12], and Fowzan S. Alkuraya, MD[1,13]

Purpose: Retinal dystrophies (RD) are heterogeneous hereditary disorders of the retina that are usually progressive in nature. The aim of this study was to clinically and molecularly characterize a large cohort of RD patients.

Methods: We have developed a next-generation sequencing assay that allows known RD genes to be sequenced simultaneously. We also performed mapping studies and exome sequencing on familial and on syndromic RD patients who tested negative on the panel.

Results: Our panel identified the likely causal mutation in >60% of the 292 RD families tested. Mapping studies on all 162 familial RD patients who tested negative on the panel identified two novel disease loci on Chr2:25,550,180-28,794,007 and Chr16:59,225,000-72,511,000. Whole-exome sequencing revealed

the likely candidate as *AGBL5* and *CDH16*, respectively. We also performed exome sequencing on negative syndromic RD cases and identified a novel homozygous truncating mutation in *GNS* in a family with the novel combination of mucopolysaccharidosis and RD. Moreover, we identified a homozygous truncating mutation in *DNAJC17* in a family with an apparently novel syndrome of retinitis pigmentosa and hypogammaglobulinemia.

Conclusion: Our study expands the clinical and allelic spectrum of known RD genes, and reveals *AGBL5*, *CDH16*, and *DNAJC17* as novel disease candidates.

Genet Med advance online publication 10 September 2015

Key Words: ciliopathy; exome; gene panel

INTRODUCTION

The human retina plays a critical role in vision by converting light energy in the visible range to electrical signals through phototransduction in photoreceptors. Photoreceptors exist in two morphologically and functionally distinct forms—cones and rods—which are respectively responsible for photopic (daylight-adapted) and scotopic (dark-adapted) vision. The term "retinal dystrophy" (RD) is used to describe a heterogeneous group of hereditary disorders in which the loss of photoreceptor function results in progressive visual impairment. The clinical variability of RD is driven primarily by the distribution of the defect. Retinitis pigmentosa, for instance, is the most common RD and is characterized by progressive loss of rods, resulting in the classical triad of reduced peripheral vision (tunnel vision), night blindness (nyctalopia), and characteristic changes in peripheral retina that can be detected by clinical examination.[1] Cone dystrophies, however, can present with reduced central vision, scotoma, photophobia, and macular lesions detectable by fundoscopy. For reasons that are not fully understood, death of one type of photoreceptor is often accompanied by death of the other type such that it can be difficult to determine the subtype of RD based on clinical retinal examination in advanced stages, which constitutes a phenotyping challenge.

The genetic heterogeneity of RD is remarkable, with 232 disease genes listed in RetNet as of February 2015 that span the

entire spectrum of Mendelian (autosomal recessive, autosomal dominant, and X-linked) and non-Mendelian (mitochondrial and, very rarely, digenic). The study of RD genetics has greatly improved our understanding of the molecular biology of vision.[2] Genes that encode components of the connecting cilium (a modified primary cilium that connects the outer and inner segments of photoreceptors) are among the most commonly mutated genes in RD, but genes encoding components of phototransductions, lipid metabolism, ion channels, extracellular matrix, retina development, intermediary metabolism, cytoskeleton, and RNA splicing are also known to be mutated in RD patients.[3] Clinically, the determination of the underlying mutation in RD is of immediate translational benefit in terms of reproductive planning and, more recently, choosing eligible patients for gene therapy for a small but growing number of RD disease genes.[4]

Despite its conspicuous basic and translational benefits, the study of RD genetics faces several important challenges. The marked locus and allelic heterogeneity of RD make it difficult to identify the causal mutation in patients, which is further compounded by the inadequate genotype/phenotype correlation.[5] Some RD disease genes are known to underlie more than one clinically distinct subtype. In addition, the current list of RD disease genes is incomplete, so a significant proportion of patients will remain undiagnosed molecularly even if they are screened for mutations in all these genes.[6] Massively parallel sequencing (also known as next-generation sequencing) has ushered in a new era in molecular diagnostics by making it feasible to sequence whole genomes, whole exomes, or large panels of selected genes. This technology is particularly suited for such a genetically and clinically heterogeneous condition as RD, and this has been demonstrated by several studies.[7–13] In this study, we describe the application of a next-generation sequencing-based gene panel strategy to study a very large cohort of RD patients in the diagnostic phase. In the discovery phase, we genotyped all familial cases that tested negative on the assay and identified three that mapped to a single locus and subjected them to whole-exome sequencing, which led us to identify *AGBL5* and *CDH16* as two novel RD candidate disease genes. In addition, we performed whole-exome sequencing on unsolved syndromic RD patients and identified *DNAJC17* as a novel candidate gene.

MATERIALS AND METHODS

Human subjects

All patients underwent ophthalmological history and clinical examination, often accompanied by electroretinography. Patients with apparently syndromic features were further evaluated by a clinical geneticist and the appropriate phenotyping tests were requested. Written informed consent was obtained from subjects before enrollment in an institutional review board (IRB)-approved protocol (KFSHRC RAC 2070023). Venous blood was collected in EDTA tubes from patients and available relatives. In selected cases, blood was also collected in sodium heparin tubes for the establishment of lymphoblastoid cell lines.

Vision loss panel assay

This assay was part of a comprehensive "Mendeliome" assay in which all established disease genes in OMIM as of August 2013 were divided into 13 clinically themed partially overlapping gene panels, which is described in detail elsewhere.[14] One of these panels is the "Vision Panel," which included 322 genes known to cause Mendelian eye diseases, including RD. The panel was used to analyze the entire study cohort. Briefly, a next-generation multiplexing assay was developed whereby the OMIM database was manually curated for genes with confirmed causal links to diseases. The resulting list of ~3,000 genes was included in 13 clinically themed panels, including the Vision Panel used in this study. Genes were amplified using the AmpliSeq HiFi mix and proprietary primer design (Thermo Fisher, Carlsbad, CA), followed by sequencing on the Ion Proton platform following the manufacturer's protocol. For the purpose of this article, we describe the results of the Vision Panel only as they pertain to the RD cohort. The variant filtering pipeline using the Saudi Human Genome is described elsewhere.[14] Briefly, only runs of high quality were aligned to the reference hg19 sequence and processed for variant calling using the Torrent Suite Variant Caller program. In the subsequent step, the variants were annotated using public-knowledge databases available in the Annovar package and other commercial data sets such as the Human Gene Mutation Database, as well as in-house databases comprising collections of disease-causing variants. Variants that are less likely to play a functional role (intronic and synonymous), those with minor allele frequency >1% and those with low base quality (using a Phred-like score produced by the Torrent Suite base caller) were excluded. The few remaining variants were then analyzed based on relevance of gene to phenotype, zygosity (when indicated), and SIFT, PolyPhen, and combined annotation-dependent depletion (CADD) scores (for missense variants). Finally, the recognized causal variant was identified as pathogenic or likely pathogenic as defined by the recent American College of Medical Genetics and Genomics guidelines.[15] All causal variants were confirmed by Sanger sequencing and segregation analysis where applicable.

Genotyping and autozygome analysis

All cases in which the vision panel failed to identify a likely causal mutation were genotyped on the Axiom SNP Chip platform according to the manufacturer's instructions (Affymetrix, Santa Clara, CA). Determination of the entire set of autozygous intervals per genome (autozygome) and autozygosity mapping were carried out as described previously.[16,17] Briefly, we used algorithms (AutoSNPa and HomozygosityMapper) that consider regions of homozygosity (ROH) >2 Mb in size as surrogates of autozygosity.

Whole-exome sequencing

All syndromic cases that tested negative on the Vision Panel, as well as negative familial cases that mapped to single loci, underwent whole-exome sequencing. One hundred nanograms of DNA sample were treated in order to obtain the

Ion Proton AmpliSeq library. Briefly, DNA was amplified in 12 separate wells using Exome Primer Pools, AmpliSeq HiFi mix (Thermo Fisher), and 10 amplification cycles. All 12 PCR pools were combined in one well and subjected to primer digestion performing incubation with FuPa reagent (Thermo Fisher). Amplified Exome targets were ligated with Ion P1 and Ion Xpress Barcode adapters after purification libraries were quantified using quantitative polymerase chain reaction with the Ion Library Quantification Kit (Thermo Fisher). The prepared exome library was further used for emulsion PCR on an Ion OneTouch System and templated Ion Sphere particles were enriched using Ion OneTouch ES; both procedures followed the manufacturer's instructions (Thermo Fisher). The template-positive Ion PI Ion Sphere particles were processed for sequencing on the Ion Proton instrument (Thermo Fisher). Approximately 15–17 Gb of sequence was generated per sequencing run. Reads were mapped to UCSC hg19 (http://genome.ucsc.edu), and variants were identified using the Saudi Human Genome Program pipeline.

RESULTS

Clinical profile

As shown in **Table 1**, a total of 541 previously unreported patients (292 families) were included in this study; 158 and 134 were multiplex and simplex families, respectively (**Supplementary Table S1** online). The majority of families were nonsyndromic ($n = 266$, 91%), and 8.9% ($n = 26$) appeared to have RD as part of a syndrome. These syndromes include Bardet-Biedl (BBS; $n = 11$), Usher ($n = 2$), and Alstrom syndromes ($n = 3$) (**Table 2**). Interestingly, one family exhibited a unique combination of skeletal features and RD unlike that in previously described oculoskeletal syndromes. Specifically, the index in 13DG0374 (**Table 2**) presented with cone-rod dystrophy in addition to short stature and narrow chest. These features overlap with those of Mainzer Saldino syndrome (caused by *IFT140* or *IFT172* mutations), which includes early-onset severe retinal dystrophy.[18,19] Spondylometaphyseal dysplasias

with cone-rod dystrophy (caused by *PCYT1A* mutations) was another differential diagnosis, but the index in 13DG0374 lacked the typical platyspondyly with anterior vertebral protrusions and progressive metaphyseal irregularity and cupping typical of that disorder.[20] Several families had other unusual syndromic associations, e.g., a family with LCA, neuroregression, and intractable seizures; one with retinitis pigmentosa and skeletal dysplasia; and another with retinitis pigmentosa, tufting enteropathy, and microcephaly (**Table 2**). In nonsyndromic cases, retinitis pigmentosa was the most common RD subtype ($n = 160$, 60.1%), followed by Leber congenital amaurosis ($n = 33$, 12.4%), cone-rod dystrophy ($n = 29$, 10.9%), and Goldmann Favre disease ($n = 13$, 4.8%) (**Table 1**).

High clinical sensitivity of the vision panel

A pathogenic or likely pathogenic causal variant was identified in 62.3% of the study cohort (182/292), 72 of which (72/182, 39.5%) are novel (**Supplementary Table S1** online). The majority of families in which a likely disease-causing mutation was identified were consistent with autosomal recessive inheritance of RD (93.4%, 170/182); these were homozygous except in ~4.4% ($n = 8/182$) of families in which compound heterozygous mutations were identified. X-linked mutations were discovered in only four families, whereas dominant heterozygous mutations were identified in eight families. This pattern is highly consistent with our experience with other genetically heterogeneous diseases in our highly consanguineous population; i.e., autosomal recessive subtypes usually dominate the genetic landscape of these diseases.[16,21] A total of 61 genes with 143 unique likely causal mutations were found in this study; their frequency distribution is shown in **Figure 1**. The clinical utility of the Vision Panel was not uniform; it was highest in conditions for which nearly all the disease genes had been identified, e.g., BBS ($n = 11$, 100%) and Alstrom syndrome ($n = 2$, 100%), but lower for nonsyndromic RD (60.5%, 161/266 cases), consistent with the notion that additional disease genes likely exist (see below).

Expanding the clinical spectrum of RD disease genes

We observed several instances of nonsyndromic RD caused by mutations in syndromic RD genes; e.g., 10DG1111 is a family with nonsyndromic rod-cone dystrophy carrying a known disease causing mutation in *BBS2*, 11DG1451 was diagnosed with nonsyndromic RP and has a known disease-causing mutation in *USH2A*, and 12DG1471 is a patient with nonsyndromic RP in whom we discovered a likely pathogenic mutation in WDR19 that is reported to cause nephronophthisis (**Supplementary Table S1** online). The nonsyndromic status of these patients was confirmed by recalling them for a full clinical assessment after their underlying mutation was revealed. One particularly surprising finding was the identification of a de novo truncating mutation in *JAG1* in a patient who presented with typical clinical findings of retinitis pigmentosa (**Supplementary Figure S1** online). Reassessment revealed no history of neonatal or infantile jaundice, congenital heart disease, short stature, or

Table 1 Overview of the clinical subtypes in the study cohort

Nonsyndromic cases	266
Retinitis pigmentosa and rod-cone dystrophy	160
Leber congenital amaurosis	33
Cone-rod dystrophy	29
Goldmann Favre disease	13
Familial exudative vitreoretinopathy with/without retinal detachment	5
Macular dystrophy	4
Stargardt disease	3
Choroideremia	2
Congenital stationary night blindness	2
Achromatopsia	1
Doyne honeycomb retinal dystrophy	1
Trichofollicular RP	1
Retinal dystrophy not otherwise specified	12

Table 2 Summary of syndromic retinal dystrophies cases

| Index ID	Phenotype	Gene name	Mutation	HGMD
08DG-00444	Retinitis pigmentosa with hypogammaglobulinemia	DNAJC17	NM_018163.3:c.681G>A; p.Tyr201_Ala227del	This study
09DG00644	Bardet-Biedl syndrome	BBS7	NM_176824.2:c.602-2A>T	CS122047
09DG01063	Bardet-Biedl syndrome	BBS7	NM_176824.2:c.602-2A>T	CS122047
10DG0923	Usher syndrome with Duane retraction syndrome	USH2A	NM_206933.2:c.4429G>T; p.Gly1477*	This study
10DG1288	Alstrom syndrome	ALMS1	NM_015120.4:c.849dupA; p.Glu284Argfs*12	This study
10DG1583	Alstrom syndrome	ALMS1	NM_015120.4:c.2729C>G; p.Ser910*	This study
11DG0880	LCA, neuroregression with global developmental delay and intractable seizures.	GUCY2D	NM_000180.3:c.2285del; p.Ser762Thrfs*22	CD1411570
12DG0327	Usher syndrome	MYO7A	NM_000260.3:c.3591_3592del; p.Cys1198Argfs*30	This study
12DG1190	Bardet-Biedl syndrome	BBS5	NM_152384.2:c.966dup; p.Ala323Cysfs*57	CI143113
12DG1296	Alstrom syndrome	ALMS1	NM_015120.4:c.1674del; p.Pro559Leufs*37	This study
12DG1870	Bardet-Biedl syndrome	BBS2	NM_031885.3:c.508G>A; p.Asp170Asn	This study
12DG2094	Bardet-Biedl syndrome	BBS10	NM_024685.3: c.1892_1896del; p.Met631Asnfs*3 & NM_024685.3:c.728_731del; p.Lys243Ilefs*15	This study and CD064500
12DG2180	Retinitis pigmentosa with hypogammaglobulinemia	NR2E3	NM_014249.3:c.373C>T; p.Arg125*	CM132550
12DG2322	Bardet-Biedl syndrome	BBS5	NM_152384.2:c.955_957del; p.Glu319del	This study
12DG2581	Retinitis pigmentosa, obesity, and mild facial dysmorphism	JAG1	NM_000214.2:c.1481dup; p.Asn494Lysfs*12	This study
12DG2641	Fundus flavimaculatus with epilepsy	RLBP1	NM_000326.4:c.773T>G; p.Leu258Trp	This study
13DG0355	MPS IIID + retinitis pigmentosa	GNS	NM_002076.3:c.83del; p.Leu28Argfs*37	This study
13DG0374	Cone-rod dystrophy, short stature, narrow chest	C21orf2	NM_001271441.1: c.1000-23A>T; .Asn334Valfs*140	This study
13DG0441	Poretti Boltshauser syndrome	LAMA1	NM_005559.3:c.7160G>A; p.Trp2387*	This study
13DG0522	Bardet-Biedl syndrome	BBS2	NM_031885.3:c.700C>T; p.Arg234*	CM051012
13DG1032	Bardet-Biedl syndrome	BBS4	NM_033028.4:c.157-2A>G	CS021722
13DG2231	Bardet-Biedl syndrome	BBS9	NM_198428.2:c.263C>A; p.Ser88*	This study
13DG2246	Bardet-Biedl syndrome	BBS4	NM_033028.4:c.1091C>A; p.Ala364Glu	CM021494
14DG0639	Severe rod-cone dysfunction, tufting enteropathy, microcephaly, and developmental delay		Unsolved	Unsolved
14DG1145	Bardet-Biedl syndrome	BBS10	NM_024685.3:c.1892_1896del;p.Met631Asnfs*3	This study
14DG1964	MMCAT (microcornea, myopic chorioretinal atrophy, and telecanthus) syndrome	ADAMTS18	NM_199355.2:c.1298C>A; p.Thr433Asn	This study

intellectual disability (he works as a biology teacher). Although he did have some facial features reminiscent of Alagille facies (primarily upslanting palpebral fissures and broad forehead) in addition to anterior segment abnormality (posterior embryotoxon and subcapsular cataract), which are also consistent with Alagille syndrome (OMIM 118450), the results of detailed cardiac, liver, and skeletal imaging and laboratory evaluation were completely normal.

We also noted several examples of phenotypic expansion for a number of genes. For example, *CACNA1F* and *GPR179* mutations have been exclusively associated with congenital stationary night blindness. However, 13DG1130 with *CACNA1F* mutation presented with cone-rod dystrophy: photophobia was the initial presentation, there was no significant nyctalopia, the disease was progressive, and ERG showed more significant photopic than scotopic involvement. Similarly, 14DG0286 with *GPR179* mutation presented with progressive night blindness,

and his eye examination revealed a classic RP picture with bone spicules, attenuated retinal arterioles, pale discs, macular edema, and cataracts. Finally, 09DG01061 with *BEST1* mutation had been diagnosed with Goldmann Favre disease, a phenotype that has not previously been associated with this gene.

Whole-exome sequencing of prescreened RD families identifies novel disease loci

We performed whole-exome sequencing on all syndromic RD cases for whom no causal mutation was identified by the Vision Panel (*n* = 5) in order to confirm that these represent novel genetic causes of a known syndrome, e.g., Poretti-Boltshauser syndrome (the causal mutation had not been published at the time of analysis) or an apparently novel syndrome. Additionally, we genotyped all familial nonsyndromic RD cases that tested negative on the panel (*n* = 59) and identified three families that mapped to single loci; therefore, they are considered potentially
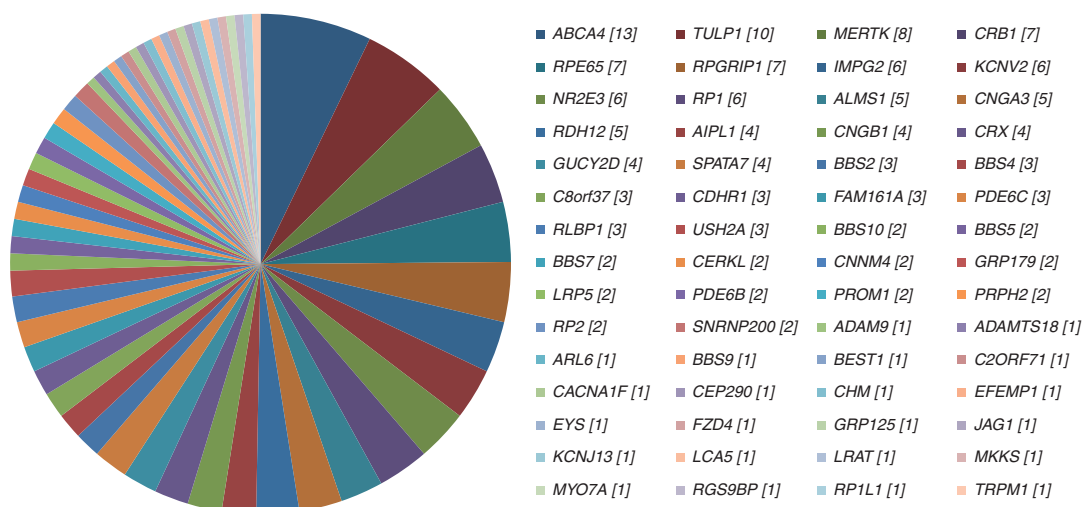
Figure 1 Distribution of the identified mutations in retinal dystrophies cases using the Vision Loss panel.

Legend entries: ABCA4 [13], TULP1 [10], MERTK [8], CRB1 [7], RPE65 [7], RPGRIP1 [7], IMPG2 [6], KCNV2 [6], NR2E3 [6], RP1 [6], ALMS1 [5], CNGA3 [5], RDH12 [5], AIPL1 [4], CNGB1 [4], CRX [4], GUCY2D [4], SPATA7 [4], BBS2 [3], BBS4 [3], C8orf37 [3], CDHR1 [3], FAM161A [3], PDE6C [3], RLBP1 [3], USH2A [3], BBS10 [2], BBS5 [2], BBS7 [2], CERKL [2], CNNM4 [2], GRP179 [2], LRP5 [2], PDE6B [2], PROM1 [2], PRPH2 [2], RP2 [2], SNRNP200 [2], ADAM9 [1], ADAMTS18 [1], ARL6 [1], BBS9 [1], BEST1 [1], C2ORF71 [1], CACNA1F [1], CEP290 [1], CHM [1], EFEMP1 [1], EYS [1], FZD4 [1], GRP125 [1], JAG1 [1], KCNJ13 [1], LCA5 [1], LRAT [1], MKKS [1], MYO7A [1], RGS9BP [1], RP1L1 [1], TRPM1 [1]

high-yield families for novel gene discovery so they were subjected to whole-exome sequencing.

Family 09DG00922 consists of five affected members who presented similarly with nonsyndromic retinitis pigmentosa. They mapped to a single locus (Chr3:94317310-105317310) that contains two known RD genes: *IMPG* and *ARL6*. Because we hypothesized that a disease-causing mutation in one of these genes may have been missed by the Vision Panel, we proceeded with Sanger sequencing, but no likely pathogenic mutation was identified at the level of genomic DNA. Therefore, we proceeded with whole-exome sequencing. However, no novel variants were identified within this locus. We considered the alternative possibility that a splicing/regulatory mutation in *IMPG* or *ARL6* was missed and preceded with reverse transcriptase–polymerase chain reaction (RT-PCR) using RNA extracted from lymphoblastoid cell lines we established for three of the affected siblings. RT-PCR did not reveal any abnormalities in *ARL6* and results were inconclusive for *IMPG2* owing to apparent lack of expression in lymphoblastoid cells. Therefore, we cannot exclude the possibility of a deep-splicing or regulatory-element mutation in *IMPG2* in this family.

Family 12DG0507 consists of three affected members with retinitis pigmentosa phenotype, which we found to map to a single novel locus, Chr2:25,550,180-28,794,007 (**Figure 2** and **Supplementary Figure S5** online). Because no known RD disease genes are present within this locus, we performed whole-exome sequencing, which revealed only one novel coding/splicing variant in the locus: a novel missense mutation in *AGBL5* (also known as *CCP5*) (NM_001035507:c.826C>T: p.Arg276Trp). This homozygous mutation is absent in 615 Saudi in-house exomes, as well as the ExAC Browser. It is predicted to be highly pathogenic in silico (PolyPhen probably damaging (0.99), SIFT deleterious (0.01)). This homozygous change fully segregated with the retinitis pigmentosa phenotype in the family (**Figure 2**). CCP5 is a tubulin-processing carboxypeptidase that is homologous to members of the M14 metallocarboxypeptidase gene family.[22] We used Phyre2 (ref. 23) to carry out

homology modeling of the human CCP5 structure (residues 156–571) with the crystal structure of Peptidase M14 (PDBID: 3L2N) as a template. The results show that p.Arg276 is located in a solvent-exposed region of the protein. The mutation of the charged Arg to a bulky hydrophobic Trp at position 276 may result in either a steric clash between the large side chain of Trp and the nearby helix or exposure of the hydrophobic Trp to the solvent (**Figure 2**). In either conformation, the p.Arg276Trp substitution potentially destabilizes the protein structure.

Family 12DG1932 consists of four affected members with late-onset RP phenotype and was found to map to a single novel locus, Chr16:59,225,000-72,511,000 (**Figure 3** and **Supplementary Figure S4** online). No known RD disease genes are present within this locus. Whole-exome sequencing revealed only one rare coding/splicing variant in the locus: a missense mutation in *CDH16* (NM_004062.3:c.950C>T: p.Ala317Val). While absent in 615 Saudi in-house exomes, this variant was present at a very low frequency in the ExAC browser (0.00006219) but never in the homozygous state. It is predicted to be pathogenic in silico (PolyPhen possibly damaging (0.83), and SIFT deleterious (0.01)). This homozygous change fully segregated with the RP phenotype in the family (**Figure 3**).

Family 08DG-00444 consists of two siblings with the apparently novel combination of retinitis pigmentosa and hypogammaglobulinemia, necessitating regular immunoglobulin infusion. Whole-exome sequencing of this syndromic form of RD revealed 83,203 variants but only one novel coding/splicing variant within the shared autozygome of the two individuals in *DNAJC17*:NM_018163:c.681G>A. Although this is a synonymous change (Ala227Ala), it affects the last base pair of exon 9, making it likely that it affects splicing by abolishing the donor site. RT-PCR using RNA extracted from lymphoblastoid cells from these two siblings confirmed the skipping of exon 9 (NM_018163:r. 601_681del; p.Tyr201_Ala227del). This homozygous variant is absent in 615 Saudi exomes and in the ExAC browser (**Supplementary Figure S2** online).
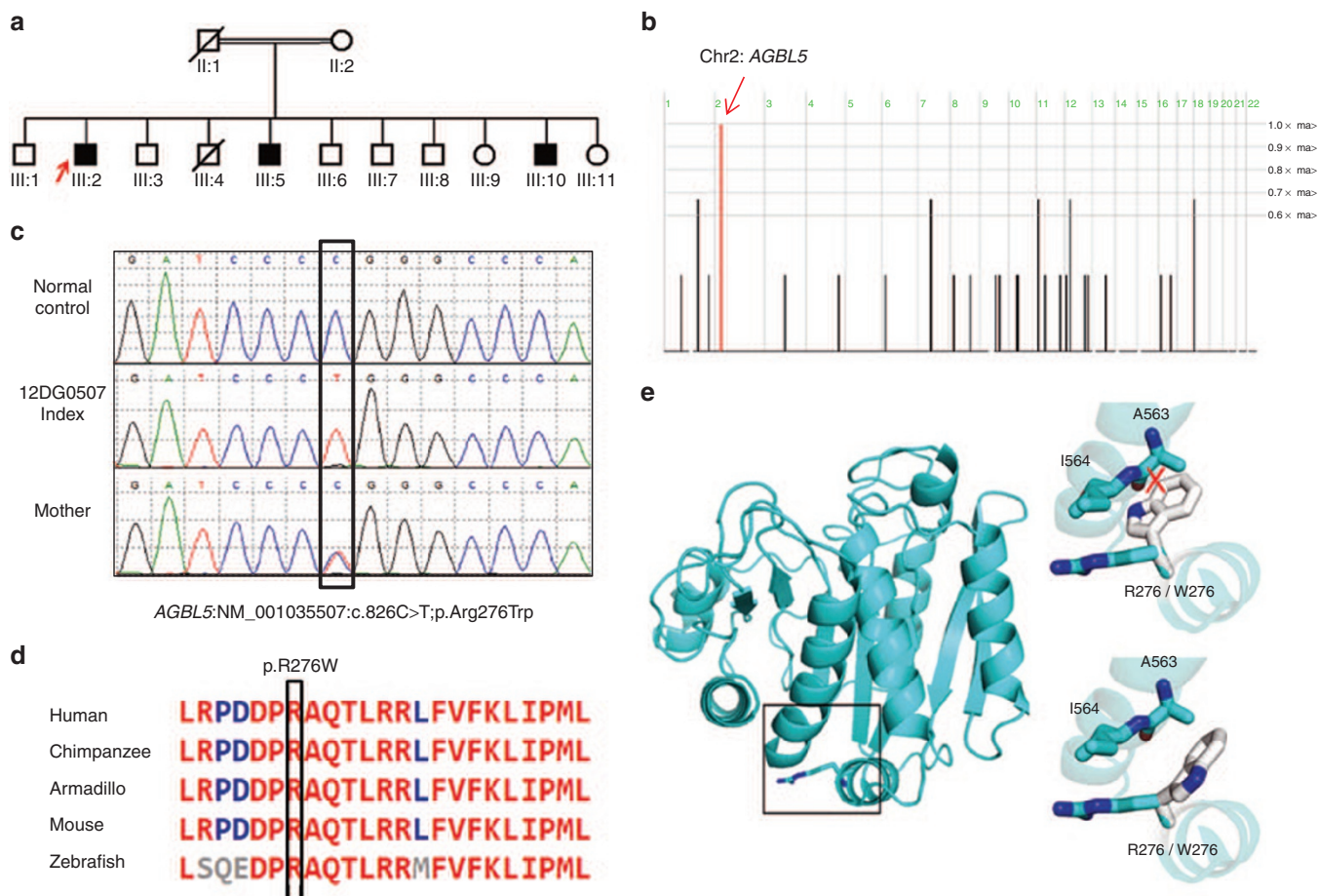
Figure 2 Identification of *AGBL5* as a novel disease candidate gene in family 12DG0507. (**a**) Pedigree for family 12DG0507. (**b**) Homozygosity Mapper utilizing three affected individuals displays a single locus mapping to *AGBL5*. (**c**) DNA sequence chromatogram of the missense mutation found in all three affected individuals. (**d**) Arg276 that is mutated in these individuals is strongly conserved. (**e**) Homology model of the catalytic domain of AGBL5. Left panel is the overall structure in ribbon representation. Arg276 (stick) is located in a solvent-exposed region. The top right panel shows Arg276Trp (Arg276 in cyan, Trp in grey) substitution potentially causes a steric clash (marked by a red "X") between the bulky Trp and a nearby helix, specifically with Ile564 and Ala563. Alternatively, as shown in the bottom right panel, Trp276 may adopt a conformation that exposes the hydrophobic side chain of Trp to the solvent.

Family 10DG1769 consists of three affected individuals who presented with short stature and cone-rod dystrophy. Whole-exome sequencing did not reveal any likely pathogenic mutation. Serendipitously, a family with Jeune asphyxiating thoracic dystrophy (*JATD*) that mapped to *C21orf2*, previously reported to cause nonsyndromic cone-dystrophy,[16] was found to have a deep intronic mutation (NM_001271441.1: c.1000-23A>T; p.Asn334Valfs*140) causing aberrant splicing (Shaheen et al., personal communication). This same mutation was found in family 13DG0374, fully segregating with the phenotype, and was absent in the rest of our in-house database of 615 whole exomes. Subsequent clinical and radiological assessments revealed subtle features of JATD (**Supplementary Figure S3** online), thus making it likely that this is a causal mutation in this family as well.

Family 13DG0355 presented with progressive skeletal deformity (dysostosis multiplex) highly consistent with the clinical suspicion of mucopolysaccharidosis (MPS), but affected members all had retinitis pigmentosa, which is not known to accompany MPS. Therefore, we thought this might represent a novel subtype of MPS. However, exome sequencing revealed a novel homozygous truncating mutation in *GNS* (**Table 2**) as the only novel coding/splicing variant within the shared autozygome. Thus, this family also represents a phenotypic expansion of a known disease gene.

Finally, family 13DG0441, which presented with Poretti-Boltshauser syndrome, was found to harbor a novel nonsense mutation in LAMA1 (NM_005559.3:c.7160G>A; p.Trp2387*), consistent with the recent discovery of this gene as the underlying cause of this very rare disorder (**Table 2**).[24]

DISCUSSION

Several strategies have been pursued to approach the genetic heterogeneity of RD. Autozygosity-guided sequencing of RD genes has been shown to identify causal mutations in the majority of RD patients in consanguineous populations, but its reliance on Sanger sequencing greatly limits its throughput.[16] In addition, its utility in outbred populations is limited. Arrays on which
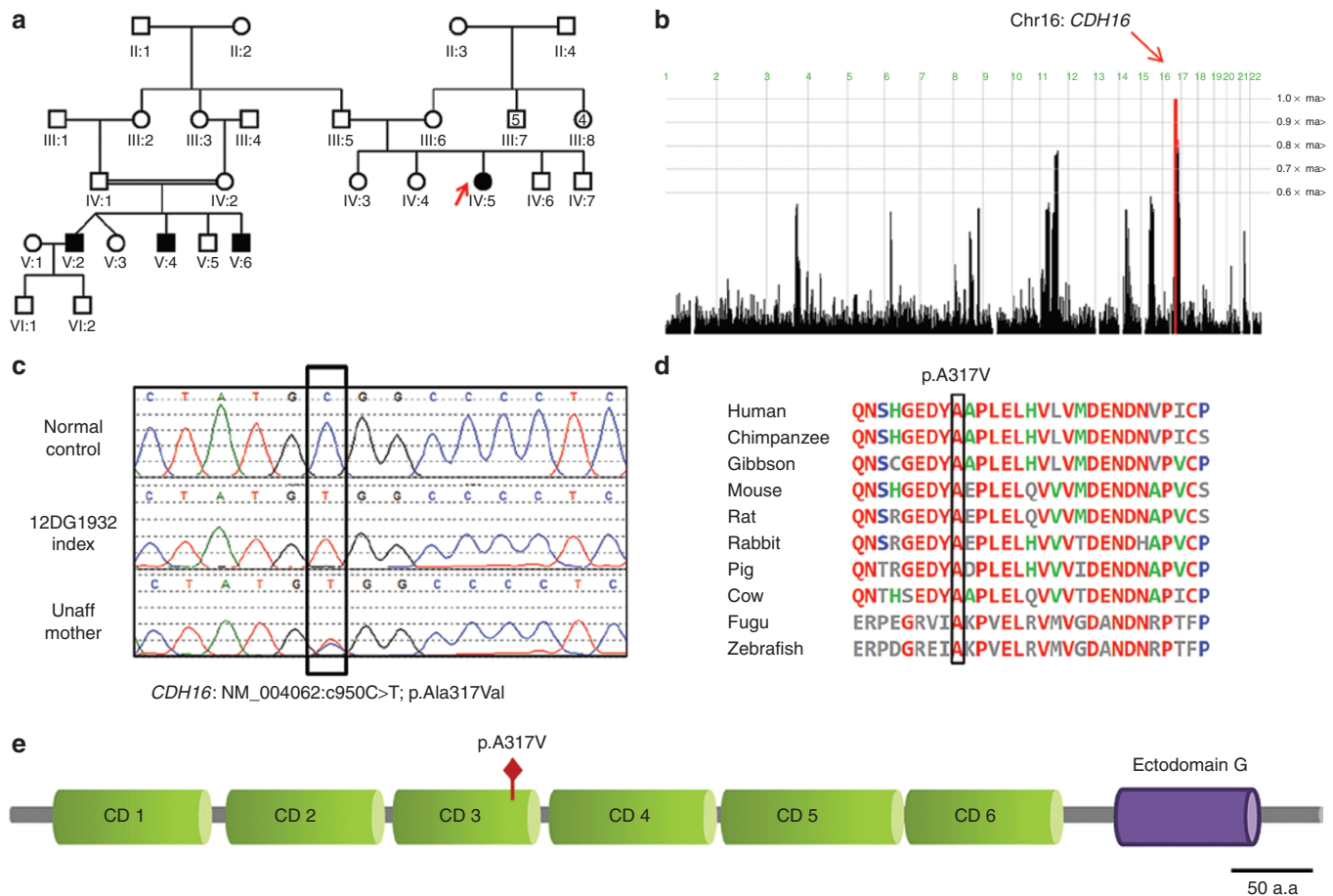
Figure 3 Identification of *CDH16* as a novel disease candidate gene in family 12DG1932. (**a**) Pedigree of family 12DG1932. (**b**) HomozygosityMapper reveals a single locus mapping to *CDH16*. (**c**) DNA sequence chromatograms of the missense mutation found in *CDH16*. (**d**) Strong conservation of Ala317 that is mutated in this family. (**e**) Protein schematic shows that the mutation affects the third cadherin domain.

previously reported RD mutations are featured cannot identify novel mutations. Next-generation sequencing offers high-throughput solutions for genetically heterogeneous conditions such as RD. Although no large-scale whole-genome sequencing studies have been performed on RD, whole-exome sequencing has been shown to be a particularly powerful diagnostic method in RD, with clinical sensitivity of up to 80% reported.[16] Whole-exome sequencing has the dual power to identify mutations in established disease genes as well as to highlight novel candidate genes. Indeed, many of the more recently described RD genes were identified through whole-exome sequencing. The relatively high cost of whole-exome sequencing and the challenge in interpreting the tens of thousands of variants that are typically generated make it attractive to consider strategies that "prescreen" cases for known RD genes before investing resources into whole-exome sequencing for the purpose of novel gene discovery. We have previously shown that prescreening RD cases with autozygome-guided Sanger sequencing can greatly accelerate the yield of novel gene discovery.[16]

Gene panels are an interesting application of next-generation sequencing that may represent another prescreening strategy. Several approaches have been taken to the design and

sequencing of gene panels in the setting of RD, and they generally resulted in a relatively high diagnostic yield of approximately 50%. In this study, we opted for a more comprehensive panel design that includes all genes known to cause Mendelian eye diseases, not only RD. The justification of this broader design is to simplify the molecular diagnostic process for the referring ophthalmologist and also to take into account the remarkable variability and allelism between many genetic eye conditions. Indeed, the adult patient with *JAG1* mutation who presented with retinitis pigmentosa and the other patients who presented with nonsyndromic RD but were found to harbor mutations in genes linked to established syndromes are good examples of such variable presentation. We speculate that this comprehensive design may have accounted, at least in part, for the exceptionally high diagnostic yield of 62.3% in our study compared with previous studies that pursued gene-panel approaches and next-generation sequencing.

The high clinical sensitivity of our assay offered us an opportunity to focus our whole-exome sequencing resources on the "negative" cases in order to increase the discovery potential. Specifically, we performed whole-exome sequencing on three multiplex nonsyndromic RD families in whom autozygosity

mapping identified single loci compatible with autosomal recessive inheritance as well as the five families that presented with syndromic RD. Interestingly, two of these syndromic RD families presented with a combination of cone-rod dystrophy and skeletal dysplasia distinct from previously reported RD/skeletal dysplasia syndromes, leading us to assume that they may have novel clinical syndromes. However, one of the two families was later found to have a deep intronic mutation in *C21orf2*, confirming that the phenotype was a variant of JATD. In our original discovery of *C21orf2* as a novel cone-rod dystrophy gene, we found that one case was truly nonsyndromic and the other was found to have short stature and skeletal abnormalities consistent with JATD.

The family with the unusual combination of MPSIIID and RP is interesting. Suzki et al. have highlighted the extremely rare association between RP and MPSII, and a more recent study has shown that the disease gene for MPSIIIC (*HGSNAT*) can cause a nonsyndromic form of RP.[25,26] It appears likely that the homozygous truncating mutation we identified in GNS is causal for both MPS and RP within this extended family, although we cannot exclude the remote possibility that it is present in tight linkage disequilibrium with another mutation causing RP.

Our above-described approach enabled us to unravel three novel candidate genes. The candidacy of *CDH16* and *AGBL5* is supported by positional mapping data because each corresponds to a single critical autozygous interval in the respective family. *AGBL5* encodes AATP/GTP-binding protein-like 5, also known as cytosolic carboxypeptidase 5 (CCP5), which is one of several cytosolic carboxypeptidases involved in the glutamylation of tubulin, a critical posttranslational modification, the disruption of which impairs microtubule-based structures, especially the primary cilium. Knockdown of ccp5 in zebrafish leads to a classic ciliopathy phenotype.[27] The biological plausibility of *CDH16*, encoding Ksp-Cadherin, is less clear, although other cadherins (PCDH15, CDH23, and CDH3) are known to cause RD. In the absence of an independent mutation in *AGBL5* and *CDH16*, their candidacy remains to be proven. Similarly, *DNAJ17*—encoding DNAJ/HSP40 homolog, subfamily C, member 11—is a gene we found to harbor a homozygous truncating mutation, and this was the only variant that survived the filtering of whole-exome sequencing in a family with the apparently novel syndrome of RD and hypogammaglobulinemia. Very little is known about this gene except that it was very recently found to harbor somatic mutations in patients with myeloproliferative disorders.[28] No functional studies were performed to follow up on the latter finding, but it does suggest a role played by this gene in bone marrow, making it possible to speculate on a potential role of the mutation we identified in the causation of hypogammaglobulinemia. Further germ-line mutations in *DNAJC17* will be needed to establish its candidacy as a potential disease gene for the syndrome of retinitis pigmentosa and hypogammaglobulinemia.

In summary, we report the largest cohort of RD (syndromic and nonsyndromic) studied by comprehensive next-generation sequencing of a gene panel and report a high diagnostic yield of >60%. This prescreening allowed us to identify RD cases in which whole-exome sequencing is likely to reveal novel candidate disease genes, *AGBL5*, *CDH16*, and *DNAJC17*, based on a single likely pathogenic homozygous mutation in each of the three genes. The majority of RD patients can now be diagnosed molecularly thanks to the advent of these new genomic tools; although novel disease genes likely exist, they will account for an increasingly small percentage of the overall patient population.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at http://www.nature.com/gim

DISCLOSURE
The authors declare no conflict of interest.

REFERENCES
1. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *Lancet* 2006;368: 1795–1809.
2. Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. *Prog Retin Eye Res* 2010;29: 335–375.
3. Wright AF, Chakarova CF, Abd El-Aziz MM, Bhattacharya SS. Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. *Nat Rev Genet* 2010;11:273–284.
4. Alkuraya FS. Saudi genetic ophthalmology research: The local and global impact. *Saudi J Ophthalmol* 2010;24:109–110.
5. Daiger SP, Sullivan LS, Bowne SJ. Genes and mutations causing retinitis pigmentosa. *Clin Genet* 2013;84:132–141.
6. Murray M, Babyatski M, Giovanni M. *Clinical Genomics: Practical Applications for Adult Patient Care*. McGraw-Hill Professional: New York, 2013.
7. Neveling K, Collin RW, Gilissen C, et al. Next-generation genetic testing for retinitis pigmentosa. *Hum Mutat* 2012;33:963–972.
8. Wang X, Wang H, Sun V, et al. Comprehensive molecular diagnosis of 179 Leber congenital amaurosis and juvenile retinitis pigmentosa patients by targeted next generation sequencing. *J Med Genet* 2013;50:674–688.
9. Shanks ME, Downes SM, Copley RR, et al. Next-generation sequencing (NGS) as a diagnostic tool for retinal degeneration reveals a much higher detection rate in early-onset disease. *Eur J Hum Genet* 2013;21:274–280.
10. Coppieters F, De Wilde B, Lefever S, et al. Massively parallel sequencing for early molecular diagnosis in Leber congenital amaurosis. *Genet Med* 2012;14: 576–585.
11. Coppieters F, Van Schil K, Bauwens M, et al. Identity-by-descent-guided mutation analysis and exome sequencing in consanguineous families reveals unusual clinical and molecular findings in retinal dystrophy. *Genet Med* 2014;16:671–680.
12. Huang XF, Huang F, Wu KC, et al. Genotype–phenotype correlation and mutation spectrum in a large cohort of patients with inherited retinal dystrophy revealed by next-generation sequencing. *Genet Med* 2015;17: 271–278.
13. Eisenberger T, Neuhaus C, Khan AO, et al. Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: the example of retinal dystrophies. *PLoS One* 2013;8:e78496.
14. Saudi Mendeliome Group. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. 2015;16:134.
15. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American

College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–424.

16. Abu-Safieh L, Alrashed M, Anazi S, et al. Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. *Genome Res* 2013;23:236–247.

17. Alkuraya FS. Discovery of rare homozygous mutations from studies of consanguineous pedigrees. *Curr Protoc Hum Genet* 2012:6.12. 11–16.12. 13.

18. Perrault I, Saunier S, Hanein S, et al. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. *Am J Hum Genet* 2012;90:864–870.

19. Halbritter J, Bizet AA, Schmidts M, et al.; UK10K Consortium. Defects in the IFT-B component IFT172 cause Jeune and Mainzer-Saldino syndromes in humans. *Am J Hum Genet* 2013;93:915–925.

20. Hoover-Fong J, Sobreira N, Jurgens J, et al. Mutations in PCYT1A, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy. *Am J Hum Genet* 2014;94:105–112.

21. Alkuraya FS. Genetics and genomic medicine in Saudi Arabia. *Mol Genet Genomic Med* 2014;2:369–378.

22. Berezniuk I, Lyons PJ, Sironi JJ, et al. Cytosolic carboxypeptidase 5 removes a- and ?-linked glutamates from tubulin. *J Biol Chem* 2013;288:30445–30453.

23. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 2009;4:363–371.

24. Aldinger KA, Mosca SJ, Tétreault M, et al.; University of Washington Center for Mendelian Genomics; Care4Rare Canada. Mutations in LAMA1 cause cerebellar dysplasia and cysts with and without retinal dystrophy. *Am J Hum Genet* 2014;95:227–234.

25. Haer-Wigman L, Newman H, Leibu R, et al. Non-syndromic retinitis pigmentosa due to mutations in the mucopolysaccharidosis type IIIC gene, heparan-alpha-glucosaminide N-acetyltransferase (HGSNAT). *Human Mol Genet* 2015: ddv118.

26. Suzuki Y, Aoyama A, Kato T, Shimozawa N, Orii T. Retinitis pigmentosa and mucopolysaccharidosis type II: an extremely attenuated phenotype. *J Inherit Metab Dis* 2009;32:582–583.

27. Pathak N, Austin-Tse CA, Liu Y, Vasilyev A, Drummond IA. Cytoplasmic carboxypeptidase 5 regulates tubulin glutamylation and zebrafish cilia formation and function. *Mol Biol Cell* 2014;25:1836–1844.

28. Hou Y, Song L, Zhu P, et al. Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell* 2012;148: 873–885.