Whole-Exome Sequencing Links a Variant in *DHDDS* to Retinitis Pigmentosa

Stephan Züchner,[1,2] Julia Dallman,[6] Rong Wen,[3] Gary Beecham,[1,2] Adam Naj,[1,2] Amjad Farooq,[4,5] Martin A. Kohli,[1,2] Patrice L. Whitehead,[1,2] William Hulme,[1,2] Ioanna Konidari,[1,2] Yvonne J.K. Edwards,[1,2] Guiqing Cai,[7] Inga Peter,[8] David Seo,[1,2] Joseph D. Buxbaum,[7,8] Jonathan L. Haines,[9] Susan Blanton,[1,2] Juan Young,[1,2] Eduardo Alfonso,[3] Jeffery M. Vance,[1,2] Byron L. Lam,[3,*] and Margaret A. Peričak-Vance[1,2,*]

Increasingly, mutations in genes causing Mendelian disease will be supported by individual and small families only; however, exome sequencing studies have thus far focused on syndromic phenotypes characterized by low locus heterogeneity. In contrast, retinitis pigmentosa (RP) is caused by >50 known genes, which still explain only half of the clinical cases. In a single, one-generation, nonsyndromic RP family, we have identified a gene, dehydrodolichol diphosphate synthase (*DHDDS*), demonstrating the power of combining whole-exome sequencing with rapid in vivo studies. DHDDS is a highly conserved essential enzyme for dolichol synthesis, permitting global *N*-linked glycosylation. Zebrafish studies showed virtually identical photoreceptor defects as observed with *N*-linked glycosylation-interfering mutations in the light-sensing protein rhodopsin. The identified Lys42Glu variant likely arose from an ancestral founder, because eight of the nine identified alleles in 27,174 control chromosomes were of confirmed Ashkenazi Jewish ethnicity. These findings demonstrate the power of exome sequencing linked to functional studies when faced with challenging study designs and, importantly, link RP to the pathways of *N*-linked glycosylation, which promise new avenues for therapeutic interventions.

Retinitis pigmentosa (RP) refers to a large group of genetically heterogeneous retinal degenerative disorders characterized by early rod photoreceptor dysfunction followed by progressive rod and cone photoreceptor dysfunction and photoreceptor death (MIM 268000). Impaired night vision followed by impaired peripheral vision generally starts in adolescence to young adulthood, with subsequent impaired central vision in later life.[1] The prevalence of RP is approximately 1 in 3000–4500 persons.[2] Autosomal-recessive, autosomal-dominant, and X-linked recessive forms of RP are all found. Mutations in over 50 genes have been identified that cause RP. These include mutations in genes that encode proteins in the phototransduction cascade, vitamin A metabolism, cellular or cytoskeletal structure, cell-to-cell signaling or synaptic interaction, RNA intron-splicing factors, and intracellular protein trafficking.[3] Genetic testing of genes known to cause RP fails to detect the disease-causing mutations in 50% of cases.[3]

We studied a family of Ashkenazi Jewish (AJ) origin in which three out of four siblings (two females and one male) were diagnosed with RP in their teenage years (Figure 1A). Early symptoms consisted of impaired night and peripheral vision. Clinical examination of the affected individuals revealed pigmentary retinal degeneration (see Figure S4 available online), and the diagnosis of retinitis pigmentosa was confirmed by impaired rod and cone responses on electroretinograms. In addition to ophthalmic evaluations, neurologic examination, X-ray bone body survey, bone density scan, echocardiogram, brain MRI with and without contrast, serum fasting cholesterol and lipid profile, serum thyroid function studies (TSH, free T4, free T3, TBG), serum insulin-like growth factor (IGF) binding protein 1, serum IGF binding protein 2, serum clotting factors (II, V, VII, VIII, IX, X, XI, XII), and antithrombin III were performed. All studies gave normal results, but individuals II:1 and II:2 had a history of lytic bone disease first diagnosed over 15 years ago. Informed consent was obtained from all individuals, and the Institutional Review Board at the University of Miami Miller School of Medicine approved the study.

To identify the genetic cause of this likely recessive subtype of RP, we screened all genes known to harbor RP mutations and found that they were negative for mutations. Classic linkage approaches were not applicable because of the size of the nonconsanguineous family, so we performed whole-exome sequencing in the three affected siblings and one unaffected sibling (Whole Human Exome Capture kit, Roche). We produced approximately 10 gigabases (Gb) of paired-end 75 bp sequence reads per individual on the Illumina GAII platform. To test the overall quality of the sequence data, we compared the genotypes of variants found in the sequence data to variants derived from genotyping via a genome-wide SNP

[1]John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL 33136, USA; [2]Dr. John T. Macdonald Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL 33136, USA; [3]Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA; [4]Department of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine, Miami, FL 33136, USA; [5]Braman Family Breast Cancer Institute, Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL 33146, USA; [6]Department of Biology, University of Miami, Miami, FL 33146, USA; [7]Department of Psychiatry, Mount Sinai School of Medicine, New York, NY 10029, USA; [8]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029, USA; [9]Center for Human Genetics Research, Vanderbilt University School of Medicine, Nashville, TN 37232, USA
*Correspondence: blam@med.miami.edu (B.L.L.), mpericak@med.miami.edu (M.A.P.-V.)
DOI 10.1016/j.ajhg.2011.01.001.

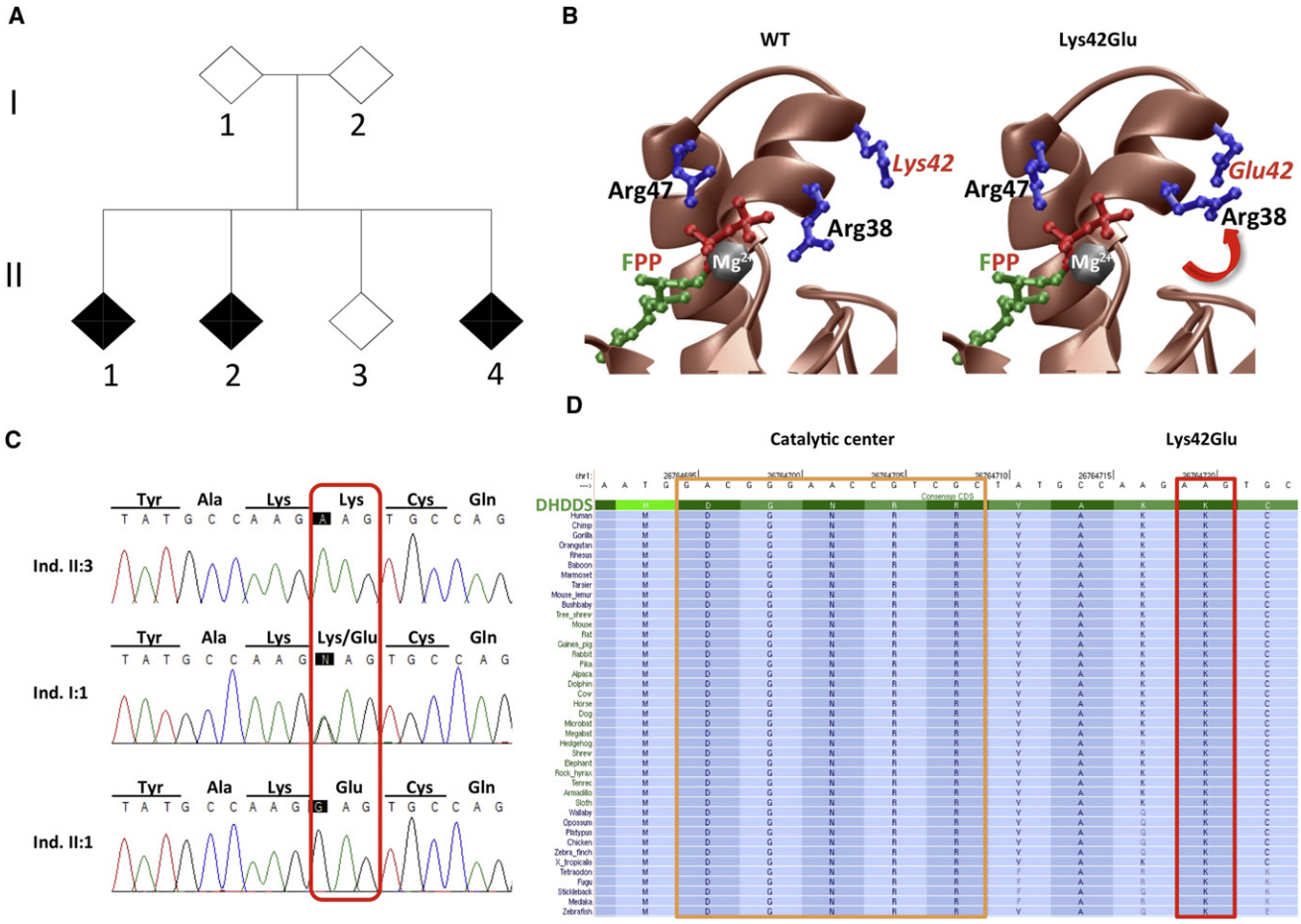

Figure 1. ***DHDDS*** **Harbors a Mutation Causing RP**
(A) Pedigree of the studied family with three affected offspring indicating a recessive trait.
(B) Three-dimensional structural model of DHDDS indicates an important role for Arg38 in the organization of the enzyme active site. The Lys42Glu mutant allele, being negatively charged, will likely ion pair with the positively charged Arg38 and, in so doing, may lead to distortion of the active site and compromise the binding of farnesyl-pyrophosphate (arrow). A detailed description of the mechanism is provided in Figure S2.
(C) Sanger sequencing confirmed the identified mutation and revealed complete cosegregation with the phenotype.
(D) Lys42 (red box) is a highly conserved residue in multiple species and is located close to the active site of the protein (orange box), indicating strong evolutionary constraints.

chip (Illumina Infinium Human 1M chip). There were 95,471 variants genotyped by next-generation sequencing at depth ≥ 8 that were also genotyped by the SNP chip. When a Phred-like quality score threshold ≥ 30 was considered, the concordance between sequencing and genotyping data was 99.7%.

Sequence reads were aligned to the human genome (hg18), and variants were called with the MAQ software package.[4] All variants were submitted to SeattleSeq for further categorization into coding, noncoding, and novel SNPs. Across the four individuals, we identified 19,307 coding and splice-site single-nucleotide variants. When filtered on homozygosity, functional significance (i.e., nonsense, missense, splice site), and single-nucleotide variants segregating in the affected, but not in the healthy, individual, 127 variants were identified (Table 1). Assuming a very rare minor allele frequency, we excluded all changes present in dbSNP131 and were left with a single variant in *DHDDS* (NM_024887.2; MIM 608172): p.Lys42Glu, c.124A>G, chr1:26637306A/G (hg18) (Table 1; Figure 1). The variant, and its cosegregation with the phenotype, was confirmed by Sanger sequencing. Both parents were heterozygous for the change, and the unaffected sibling inherited two reference alleles. Of the 231,533 raw indel counts across four exomes, 27 segregated with disease status after quality measures and 0 were novel in dbSNP131.

To test the population frequency, we genotyped the chr1:26637306A/G variant in 13,587 controls via a custom TaqMan SNP genotyping assay. The studied family was of AJ ethnicity by self-report, which was confirmed through clustering with known AJ samples in a principal component analysis of the genotype data.[5] As such, we divided the controls into three groups: (1) a group of 717 "AJ" white controls that was made up of 220 individuals of self-reported AJ ancestry and 497 individuals that clustered with

Table 1. Number of Identified Single-Nucleotide Variants Compared to the Reference Genome hg18 and Their Categorization and Filtering

	Affected Sibling 1	Affected Sibling 2	Affected Sibling 3	Shared between Affected Siblings 1 and 2	Shared between Affected Siblings 1, 2, and 3	+ NOT in Unaffected Sibling 4
Missense, nonsense, splice-site variants	8712	8716	8752	5289	4507	ND
+ Homozygous state	2516	2489	2649	1411	981	127
+ Not reported in dbSNP131	11	18	27	5	4	1

ND denotes not determined.

known AJ in a principal component analysis based on genome-wide genotype data, (2) a group of 6977 "non-AJ" white controls that did not cluster with the known AJ samples, and (3) a group of 5893 white controls without genome-wide genotype data. The chr1:26637306A/G variant was observed only eight times among the 1434 chromosomes from the AJ controls. Each time, the variant was observed as a heterozygote (minor allele frequency [MAF] = 0.0056). This confirms that the variant is indeed rare among the AJ population and that homozygotes should be very rare (frequency = 3×10^{-5}, or 3 in 100,000, assuming Hardy-Weinberg equilibrium). The variant is extremely rare in non-AJ populations; it was not observed in any of the confirmed non-AJ controls (0 out of 13,954 chromosomes; MAF < 0.00008) and was observed only once in controls without genome-wide genotype data (1 out of 11,786 chromosomes; MAF < 0.00009) (Table S1).

The array-based genotype data were also utilized post hoc to attest for consistency with the hypothesis that this mutation is causal for RP. Two different, though related, methods were applied. The first was to search for runs of homozygosity present among the affected siblings but not present in the remaining family members. This was accomplished with the PLINK software.[6] To determine large blocks of homozygosity, we required "homozygous" segments to be at least 500 kb long and have at least 95% homozygosity. There were 91 such regions that were homozygous among all three affected siblings and were not homozygous among the parents and unaffected sibling (Table S3). One of these regions on chr1:25.7Mb–chr1:29Mb contained *DDHDS*. The second method to prioritize variants was to examine estimated identity by descent (IBD) to determine regions for which all affected siblings share the same maternal and paternal haplotypes (regardless of homozygosity or heterozygosity). There were 33 regions estimated to be completely shared by all three affected siblings. *DHDDS* falls in one of these regions. We identified four known RP genes within the homozygosity and IBD sharing regions (*PDE6B* [MIM 180072], *USH1G*/*SANS* [MIM 706696], *PRCD* [MIM 610598], and *CA4* [MIM 114760]).[7] Mutations in these genes were excluded by Sanger sequencing (*PDE6B*, *USH1G*/*SANS*, *CA4*) and/or exome sequencing (*PRCD*).

DHDDS is a key enzyme in the final steps of dolichol-pyrophosphate synthesis.[8] As shown in Figure 1D, the Lys42 residue is highly conserved in different species and is located close to the catalytic center and the substrate binding site for farnesyl pyrophosphate phosphate (FPP) of the enzyme.[9] The observed Lys>Glu amino acid change at position 42 is predicted to be "damaging" by Polyphen[10] and received a high 5.25 genomic evolutionary rate profiling conservation score.[11] To further understand the physiological significance of the Lys42Glu mutation, we applied the MODELER software package[12] to generate 3D atomic models of human DHDDS in complex with FPP/Mg2+ on the basis of amino acid sequence ortholog of *E. coli* dhdds, for which there is a known crystal structure in complex with FPP/Mg2+[13] (Figure 1B and Figure S1). Within the enzyme active site, the farnesyl tail of FPP inserts deep into a hydrophobic cavity stabilized by numerous intermolecular van der Waals contacts, with the basic residues Arg38 and Arg47 playing a key role (Figure 1). Residue Lys42 is sterically close to residue Arg38 and facilitates its optimal orientation for ion pairing with the pyrophosphate group of FPP via charge-charge repulsive forces. Remarkably, the mutant residue Glu42 will compete with the pyrophosphate group of FPP for ion pairing with Arg38, resulting in the misorientation of guanidine moiety of Arg38. This suggests a loss-of-function mechanism in effectively binding the substrate FPP.

To determine whether insufficient DHDDS activity induces photoreceptor degeneration, we knocked down the expression of the highly conserved *dhdds* (NM_213187.1) in zebrafish by injecting a splice-junction blocking morpholino (MO) into fertilized eggs. MO oligonucleotides (GeneTools LLC) designed against exon1/intron1 of the zebrafish *dhdds* were injected into one-cell embryos (Figure S2). Embryos were allowed to develop for 4 days. At this age, they reliably respond to abrupt changes in light intensity with bursts of swimming (Movies S1–S3). *Dhdds* morphants exhibited a range of phenotypes typical of such knockdown experiments. In functional assays, MO-injected fish failed to react to light on-off switches in contrast to controls that reacted to light on-off switches with typical escape responses (see Figure S5 and Movies S1–S3). Upon microscopic examination, the gross pathology of the retina was intact in both the wild-type and the MO-treated animals (Figure 2). However, the photoreceptor outer segments

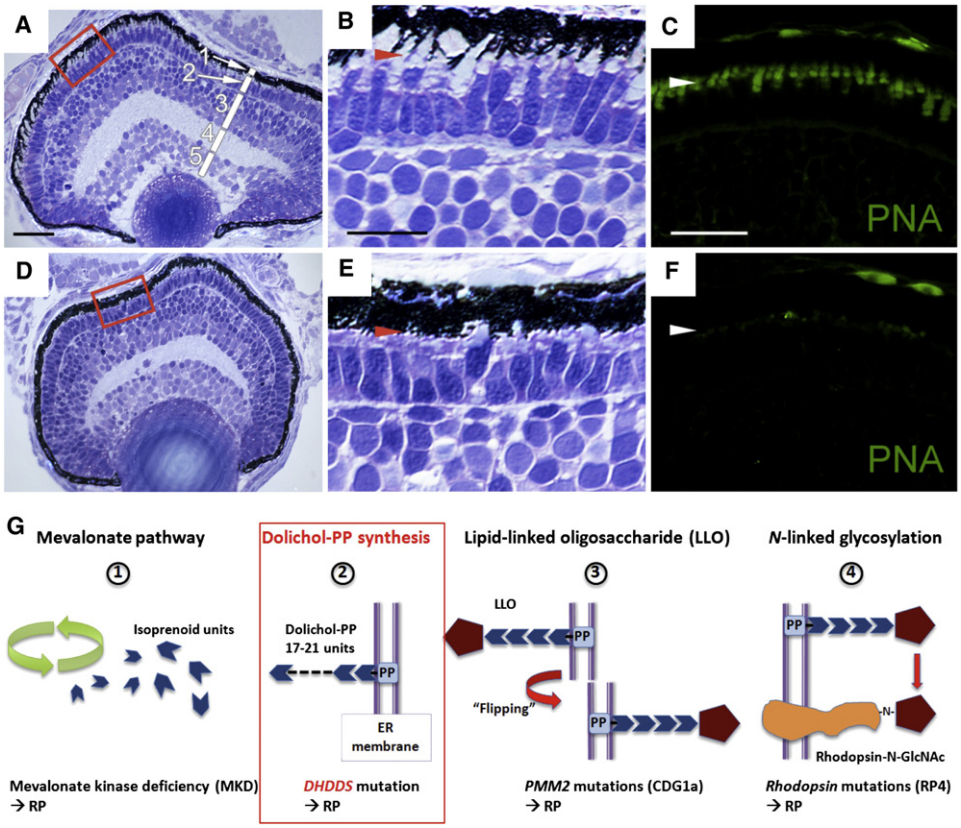

Figure 2. Photoreceptor Degeneration Induced by Suppression of *DHDDS* Expression and Illustration of the Involved Pathways of *N*-Linked Glycosylation

(A–F) Microphotographs in (A) and (D) show representative retinal sections of control and fish treated with MO. The areas in red rectangles are shown at higher magnification in (B) and (E). The layers of the retina are indicated by the white bars in (A), including the retinal pigment epithelium (1), the outer nuclear layer (2), the inner nuclear layer (3), the inner plexiform layer (4), and the retinal ganglion cell layer (5). In the retinas treated with MO, the outer segments of photoreceptors are very short, or completely missing (E, arrowhead), compared to the controls (B, arrowhead). Staining of cone outer segments with peanut agglutinin (PNA) highlights that cone outer segments in MO-treated zebrafish (F, arrowhead), but not in controls (C, arrowhead), are missing. Semithin plastic sections (A, B, D, and E) were stained with toluidine blue. Cryosections (C and F) were stained with PNA (conjugated with Alexa Fluor 488). Scale bars represent 25 μm for (A) and (D), 10 μm for (B) and (E), and 20 μm for (C) and (F).

(G) A simplified schematic of essential *N*-linked glycosylation pathways (steps 1–4). An emphasis is given to key processes upstream and downstream of dolichol-PP synthesis (DHDDS) that are also related to RP phenotypes. (1) The cytoplasmic mevalonate pathway produces isoprenoide units, which are (2) chained to ER membrane-bound dolichol species, dolichol-pyrophosphate (dolichol-PP). (3) Dolichol-PP serves as an intermediate lipid to bind a common core of carbohydrates, en bloc, forming the lipid-linked oligosaccharide (LLO). The LLO is "flipped" within the ER membrane to face the ER lumen. (4) Finally, the common core of carbohydrates is transferred to proteins, such as rhodopsin. Along this pathway, mevalonate kinase deficiency is associated with RP,[14] mutations in *PMM2* (phosphomannomutase 2; MIM 601785) lead to congenital disorder of glycosylation Ia (CDGIa), often involving RP, and, finally, glycosylation defects in rhodopsin cause RP. N denotes asparagine.

were very short or completely missing in the retinas treated with *dhdds* morpholinos (Figure 2; Figure S3). In addition to a reduced lights-off response and reduced or absent photoreceptor outer segments, one-third of morphants had smaller eyes, with a slight ventral flexion to the body axis; this indicates that, depending on degree of *dhdds* knockdown, a broader set of tissues are affected, with the most sensitized being photoreceptors. These results are consistent with *dhdds* being a key protein for the generation and continuous renewal of photoreceptor outer segments.

What is the functional role of DHDDS? About 65% of all mammalian proteins contain at least one *N*-linked glycosylation site (Asn-X-Ser(Thr)). In generating *N*-linked glycoproteins, first a common core of carbohydrate is attached to a lipid intermediate that resides in the endoplasmic reticulum membrane: dolichol-pyrophosphate (dolichol-PP). Only after this lipid-linked oligosaccharide structure has formed can the transfer of oligosaccharides to proteins occur to form glycoproteins. DHDDS catalyzes the formation of dolichol-PP by chaining 17–21 isoprene units. These isoprene units are a product of the mevalonate pathway. Thus, DHDDS connects the upstream mevalonate pathway to the membrane-bound *N*-linked glycosylation machinery supported by a multitude of enzymes and cofactors. As illustrated in Figure 2G, along these pathways, several clinically complex and heterogeneous genetic conditions involve RP as a symptom: (1) mevalonate kinase deficiency (MIM

610377) can present with RP,[14] (2) congenital disorder of glycosylation Ia (MIM 212065) is frequently associated with RP,[15] and (3) mutations in glycosylation sites of the light-activated G protein-coupled receptor rhodopsin (MIM 180380), in itself a cause for RP,[16] lead to shortening of rod outer segments in the retina of *X. laevis*.[17] These latter changes are virtually identical to those we observed in zebrafish (Figure 2). Rhodopsin is responsible for light capture and resides in the outer segment of photoreceptor cells. Given that the outer segment, together with the rhodopsin it contains, is continuously being renewed,[18] it is conceivable that this highly specialized structure has a pronounced susceptibility to defects of *N*-linked glycosylation, as would be caused by less-efficient production of dolichols. We therefore suggest that the identified mutation causes reduced, rather than complete, loss of enzymatic function of DHDDS. Notably, two of the affected subjects were diagnosed with lytic bones, potentially indicating additional subclinical consequences of the DHDDS defect, although the literature does not support a clear connection of lytic bones with diseases of glycosylation. A reduced enzymatic activity of DHDDS, due to insufficient substrate interaction, might eventually provide an opportunity for a therapeutic intervention.

In summary, we have shown that even under most constrained preconditions, such as a single nuclear family with only three affected siblings suffering from a genetically highly heterogeneous "pure" phenotype, a causative variant linked to RP can be identified by combining human genomic sequencing approaches with rapid animal modeling and in silico prediction of protein function. It thus appears to be possible to overcome some of the interpretation challenges that increasingly confront Mendelian genetics and genomic medicine.

Supplemental Data

Supplemental Data include five figures, three tables, and three movies and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

We are thankful to the family members studied and to their support for our research. We also thank Olaf Bodamer for helpful discussions and Yiwen Li, Deqiang Huang, and Zhengying Wang for technical assistance on the development of this manuscript. This study was supported by grants from the Department of Defense (W81XWH-09-1-0674), National Institutes of Health (P30-EY14801 core grant, R01-EY018586 to R.W., R01-EY012118 to M.A.P.-V., R01-GM083897 to A.F., and U54-NS065712 to S.Z.), Hope for Vision, an unrestricted grant from Research to Prevent Blindness, and a grant from the Florida Office of Tourism, Trade and Economic Development.

Received: October 14, 2010
Revised: January 4, 2011
Accepted: January 10, 2011
Published online: February 3, 2011

Web Resources

The URLs for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

SeattleSeq Annotation, http://gvs.gs.washington.edu/SeattleSeqAnnotation/

References

1. Heckenlively, J.R., Yoser, S.L., Friedman, L.H., and Oversier, J.J. (1988). Clinical findings and common symptoms in retinitis pigmentosa. Am. J. Ophthalmol. *105*, 504–511.
2. Pagon, R.A. (1988). Retinitis pigmentosa. Surv. Ophthalmol. *33*, 137–177.
3. Hartong, D.T., Berson, E.L., and Dryja, T.P. (2006). Retinitis pigmentosa. Lancet *368*, 1795–1809.
4. Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. *18*, 1851–1858.
5. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.
6. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
7. Goodwin, P. (2008). Hereditary retinal disease. Curr. Opin. Ophthalmol. *19*, 255–262.
8. Endo, S., Zhang, Y.W., Takahashi, S., and Koyama, T. (2003). Identification of human dehydrodolichyl diphosphate synthase gene. Biochim. Biophys. Acta *1625*, 291–295.
9. Fujihashi, M., Zhang, Y.W., Higuchi, Y., Li, X.Y., Koyama, T., and Miki, K. (2001). Crystal structure of cis-prenyl chain elongating enzyme, undecaprenyl diphosphate synthase. Proc. Natl. Acad. Sci. USA *98*, 4337–4342.
10. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.
11. Cooper, G.M., Stone, E.A., Asimenos, G., Green, E.D., Batzoglou, S., and Sidow, A.; NISC Comparative Sequencing Program. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. *15*, 901–913.
12. Martí-Renom, M.A., Stuart, A.C., Fiser, A., Sánchez, R., Melo, F., and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. *29*, 291–325.
13. Guo, R.T., Ko, T.P., Chen, A.P., Kuo, C.J., Wang, A.H., and Liang, P.H. (2005). Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: Roles of the metal ion and conserved residues in catalysis. J. Biol. Chem. *280*, 20762–20774.
14. Balgobind, B., Wittebol-Post, D., and Frenkel, J. (2005). Retinitis pigmentosa in mevalonate kinase deficiency. J. Inherit. Metab. Dis. *28*, 1143–1145.
15. Grünewald, S. (2009). The clinical spectrum of phosphomannomutase 2 deficiency (CDG-Ia). Biochim. Biophys. Acta *1792*, 827–834.

16. Sullivan, L.S., Heckenlively, J.R., Bowne, S.J., Zuo, J., Hide, W.A., Gal, A., Denton, M., Inglehearn, C.F., Blanton, S.H., and Daiger, S.P. (1999). Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa. Nat. Genet. *22*, 255–259.
17. Tam, B.M., and Moritz, O.L. (2009). The role of rhodopsin glycosylation in protein folding, trafficking, and light-sensitive retinal degeneration. J. Neurosci. *29*, 15145–15154.
18. Young, R.W. (1967). The renewal of photoreceptor cell outer segments. J. Cell Biol. *33*, 61–72.