Predicting the Pathogenicity of *RPE65* Mutations

A.R. Philp[1A,1B,#], **M. Jin**[2A,#], **S. Li**[2A], **E.I. Schindler**[1A], **A. Iannaccone**[3], **B.L. Lam**[4], **R.G. Weleber**[5], **G.A. Fishman**[6], **S.G. Jacobson**[7], **R.F. Mullins**[1A], **G.H. Travis**[2A,2B], and **E.M. Stone**[1A,1B]

[1A]Department of Ophthalmology and Visual Sciences, University of Iowa Hospitals and Clinics, Iowa City

[1B]the Howard Hughes Medical Institute, University of Iowa Hospitals and Clinics, Iowa City

[2A]Jules Stein Eye Institute, University of California, Los Angeles, School of Medicine, Los Angeles, California

[2B]the Department of Biological Chemistry, University of California, Los Angeles, School of Medicine, Los Angeles, California

[3]the Retinal Degeneration and Ophthalmic Genetics Service, Hamilton Eye Institute, Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, Tennessee

[4]the Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, Florida

[5]the Macular Degeneration Center, Casey Eye Institute, Oregon Health and Science University, Portland, Oregon

[6]the Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, Illinois

[7]the Scheie Eye Institute, Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania.

Abstract

To assist in distinguishing disease-causing mutations from non-pathogenic polymorphisms, we developed an objective algorithm to calculate an "estimate of pathogenic probability" (EPP) based on the prevalence of a specific variation, its segregation within families, and its predicted effects on protein structure. Eleven missense variations in the *RPE65* gene were evaluated in patients with Leber congenital amaurosis (LCA) using the EPP algorithm. The accuracy of the EPP algorithm was evaluated using a cell-culture assay of RPE65-isomerase activity The variations were engineered into plasmids containing a human *RPE65* cDNA and the retinoid isomerase activity of each variant was determined in cultured cells. The EPP algorithm predicted eight substitution mutations to be disease-causing variants. The isomerase catalytic activities of these *RPE65* variants were all less than 6% of wild-type. In contrast, the EPP algorithm predicted the other three substitutions to be non-disease-causing, with isomerase activities of 68%, 127% and 110% of wild-type, respectively. We observed complete concordance between the predicted pathogenicities of missense variations in the *RPE65* gene and retinoid isomerase activities measured in a functional assay. These results suggest that the EPP algorithm may be useful to evaluate the pathogenicity of missense variations in other disease genes where functional assays are not available.

Corresponding authors: Edwin M. Stone, M.D., Ph.D. Department of Ophthalmology and Visual Sciences The University of Iowa 375 Newton Road 4111 MERF Iowa City, Iowa 52242 E-mail: edwin-stone@uiowa.edu Phone: 319-335-8270 Fax: 319-335-7142 Gabriel H. Travis, MD Jules Stein Eye Institute at UCLA E-mail: travis@jsei.ucla.edu Phone: 310-267-2673 .
[#]Contributed equally to the work.

Keywords

Leber congenital amaurosis; pathogenicity; RPE65; retinoid

INTRODUCTION

By themselves, DNA sequence variations are often not very useful for guiding patient care. The relatively rare disease-causing variants must be distinguished from the thousands of non-disease-causing polymorphisms that exist in the genome of every individual. For recessive diseases, it is relatively straightforward to predict the pathogenicity of deletions, nonsense mutations, splice-site alterations and frameshift mutations, where the encoded protein is expected to be truncated or absent. However, the pathogenicity of missense mutations, especially those observed only rarely, can be more challenging to predict. In addition, specific variants (both disease-causing and non-disease-causing) occur at different frequencies in different ethnic groups. Ideally, functional assays could be used to directly assay the effect of each newly identified sequence variant. However, the functions of many proteins encoded by disease genes are poorly understood. Of those proteins that have been characterized functionally, many are difficult to assay *in vitro*. Finally, many disease-causing genes, including *RPE65* (MIM# 180069) exhibit a large number of different variants in large, ethnically complex, outbred populations, like those found in North America (see http://www.retina-international.org/sci-news/rpe65mut.htm). In combination, these factors make it impractical to perform a functional assessment of every genetic variant suspected to cause disease.

Recognizing the need for an objective means of using readily available data to determine the likelihood that a given mutation causes disease, our laboratory previously described an empiric algorithm known as the estimate of pathogenic probability (EPP)(Stone, 2003). The EPP takes into consideration the relative prevalence of a variation in affected individuals versus ethnically matched controls, the probable effect of the variation on the protein (using the blosum 62 substitution matrix) (Henikoff and Henikoff, 1992) and the segregation of the variation within families(Stone, 2003). An EPP of zero is assigned to variants that are unlikely to cause disease. Such variants are often found with nearly equal frequency in patients and ethnically matched controls, do not segregate with the disease phenotype in affected families, and are not predicted to alter protein function to a significant degree. An EPP of three is assigned to variants that are very likely to cause disease. Such variants are often 100-fold more common in patients than in ethnically matched controls, segregate perfectly with the disease phenotype in all families in which they are observed, and are predicted to alter the structure of the encoded protein in a biologically meaningful way. Epidemiologic and functional data have also recently been integrated into the investigation of other diseases, including hereditary colon cancer and breast cancer (Barnetson, et al., 2008; Goldgar, et al., 2008)

To validate the ability of the EPP algorithm to predict pathogenicity, we tested its predictions for selected *RPE65* variants using a functional assay for RPE65 activity, similar to assays used in previous reports of cultured cells and *Rpe65*-deficient mice(Bereta, et al., 2008; Takahashi, et al., 2006). In 1997, mutations in the *RPE65* gene were shown to cause the severe blinding-disease of childhood, Leber congenital amaurosis (LCA) in a subset of cases(Gu, et al., 1997; Marlhens, et al., 1997). Since its initial discovery, over 60 different pathogenic alterations in the RPE65 gene have been described. The first step in light perception is the absorption of a photon by a photopigment molecule in a rod or cone photoreceptor cell. This induces photoisomerization of the constituent 11-*cis*-retinaldehyde (11-*cis*-RAL) chromophore. The resulting all-*trans*-retinaldehyde (all-*trans*-RAL)

dissociates from the bleached pigment, rendering it insensitive to light. Restoration of light-sensitivity requires chemical re-isomerization of the all-*trans*-RAL back to 11-*cis*-RAL via a multi-step enzyme pathway called the visual cycle. Recent studies have shown that RPE65 functions as the retinoid isomerase in the visual cycle (Jin, et al., 2005; Moiseyev, et al., 2005; Redmond, et al., 2005). Patients with *RPE65*-mediated LCA are therefore expected to have greatly reduced or absent retinoid-isomerase activity.

MATERIALS AND METHODS

Selection of Missense Variants and Calculation of EPP

We recently screened a cohort of more than 600 LCA patients for variations in *RPE65* and seven other genes known to cause this disease (Stone, 2007). We selected 11 missense variations in *RPE65* for further study. These variants belonged to three categories: (1) variants predicted to alter protein function that were observed in multiple affected individuals in the large cohort, (2) novel variants for which no segregation data were available, and (3) variants previously suggested to cause disease in the literature but which the EPP algorithm predicted to be non-disease-causing. The EPP value for each variant was calculated as previously described (Stone, 2003). Briefly, the EPP is determined by combining information from functional assays, structural analysis, published association data and the average evolutionary effect of amino acid substitutions as predicted by the blosum 62 substitution matrix(Henikoff and Henikoff, 1992). Disease-causing variants are expected to be more common in patients than controls and to segregate correctly in families (e.g., unaffected siblings should not share genotypes with affected siblings). For autosomal recessive diseases like LCA, when a given variation is observed frequently enough that it can be shown to be more than 100 times rarer in the control group than the disease group, it receives an additional "point" in favor of pathogenicity. Similarly, when the frequency of an allele in the control population is more common than one would predict for a highly penetrant recessive allele (using the Hardy-Weinberg equation and the known frequency of the disease in the population) a point is deducted from the EPP.

Construction of the *RPE65* Expression Vector and Site-Directed Mutagenesis

The complete coding region of human *Rpe65* cDNA (RefSeq NM_000329.2) was subcloned into the mammalian expression vector, pRK5 (BD Pharmingen). Specific sequence variations were introduced using a site-directed mutagenesis kit, QuickChangeXL (Stratagene, La Jolla, CA). All plasmids were bi-directionally sequenced following mutagenesis to confirm that the expected mutations had been introduced and that the constructs were not otherwise altered. Plasmids for transfection were purified using the PureLink Hipure plasmid miniprep kit (Invitrogen, Carlsbad, CA).

Cell Culture and Transfection

HEK293T-derived 293T-LC cells that stably express lecithin:retinol acyl transferase (LRAT; MIM# 604863) and cellular retinaldehyde-binding protein (CRALBP; MIM# 180090) have been previously described (Jin, et al., 2005). These cells were grown in D-MEM (Invitrogen, Carlsbad, CA) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and antibiotics (100 units/ml penicillin G and 100 μg/ml streptomycin) at 37°C under 5% CO_2. Transfection of the cells with plasmid DNA was performed using PolyFect transfection reagent (Qiagen, Valencia, CA) according to the manufacturer's instructions.

Isomerohydrolase Assay

To assay the ability of each variant to catalyze 11-*cis*-retinol (11-*cis*-ROL) synthesis from all-*trans*-retinol (all-*trans*-ROL) substrate, 293T-LC cells were grown in 12-well culture

plates and transfected with one of the Rpe65 expression plasmids. Thirty hours after transfection, cell media were replaced with fresh media containing 15% FBS, 0.5% BSA and 25 mM HEPES buffer (pH 7.5). All-*trans*-ROL dissolved in ethanol was added to the medium (5 μM final concentration) under dim red light and incubated with the cells for additional 4, 6 or 18 hours. After removing the media, the cells were washed in 1.0 ml of PBS, pelleted by low-speed centrifugation, and lysed in 0.4 ml lysis buffer (0.1% SDS in 10 mM HEPES buffer, pH 7.5). Proteins in the cell lysates (10 μg) were analyzed by immunoblot using *RPE65* antiserum (see below). For high performance liquid chromatography (HPLC) analysis, retinoids in the cell lysates were extracted with hexane and analyzed as described previously (Jin, et al., 2007; Mata, et al., 2002; Moiseyev, et al., 2005). Briefly, samples were dissolved in hexane and the retinoids were separated on a silica column (Supelxosil LS-SI 5 mm, 4.6 × 250 mm ID) using a gradient elution 0.2-10% dioxane in hexane at 2.0 ml per minute flow rate. Identified peaks were confirmed by spectral analysis and co-elution with authentic retinoid standards. Enzymatic activity was normalized to that of a wild-type *RPE65* construct.

Immunoblot Analysis

Proteins in Laemmli sample buffer were heated to 75°C for 5 minutes, separated by SDS-PAGE in a 12% polyacrylamide gel, and transferred to an Immobilon-P membrane (Millipore). The membrane was incubated in blocking buffer (pH 7.4 PBS, 5% non-fat milk) for two hours at 37°C, then overnight at 4°C with a rabbit polyclonal antibody directed against the peptide NFITKINPETLETIK (residues #150-163 in RPE65) (Mata, et al., 2004). After washing in PBS for 30 minutes the membrane was incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Jackson ImmunoResearch Labs) or with IRDye 800CW-conjugated goat anti-rabbit IgG (LI-COR Biosciences) for one hour and washed again. The immunoreactions were visualized with the enhanced ECL-Plus Western blot detection system (GE Healthcare/Amersham), or were quantified by scanning the membrane in an Odyssey Infrared Imaging System (LI-COR Biosciences) using 780-nm excitation and 800-nm detection wavelengths. The fluorescence intensity of RPE65 band was measured using Odyssey 2.1 software (LI-COR Biosciences) (Jin, et al., 2007).

Real-time RT-PCR analysis

Messenger RNA levels of normal and mutated Rpe65 in the transfected cells were analyzed by quantitative real-time RT-PCR using a human Rpe65-specific primer pair (5′-GAACTGTCCTCGCCGCTCAC and 5′-GCAGGGATCTGGGAAAGCAC) and the SuperScript III Platinum CellsDirect Two-Step qRT-PCR kit (Invitrogen). The RT-PCR reaction was performed on a DNA Engine Opticon2 (MJ Research) according to the manufacturer's instructions. Fluorescence signals produced by binding of SYBR Green to new double-stranded amplicons were collected after each PCR cycle. Amplicon size and reaction specificity were confirmed by electrophoresis on a 2.5% agarose gel. Each Rpe65 construct was transfected into the cells in triplicate and the target mRNA levels in each plate of transfected cells were analyzed in duplicate. To normalize template input, human *GAPDH* mRNA level was measured for each samples using the human *GAPDH*-specific primer pair (5′-GGAAGGTGAAGGTCGGAGTCA; 5′-CTTCCCGTTCTCAGCCTTGAC). Relative content of *RPE65* mRNA was calculated based on its C_t (threshold cycle) relative to that of *GAPDH*. The average mRNA levels of mutated *RPE65* were compared to those of the wild-type control.

RESULTS

Molecular screening and EPP analysis

In a previous study we screened 642 patients diagnosed with LCA and 200 control individuals for variations in *RPE65*, and identified 38 different alleles that were plausibly disease-causing (i.e., an EPP of 2 or 3), and 14 alleles that were likely to be non-pathogenic (an EPP 0 or 1). In the present study, we investigated 11 putative LCA-causing alleles, divided into three groups. The first group consisted of five missense variations (p.Arg91Trp, p.Tyr239Asp, p.Gly40Ser, p.Arg44Gln and p.Arg91Gln) that had been observed in multiple unrelated patients with LCA, but not in control subjects. The second group consisted of three variations (p.Thr101Ile, p.Tyr318Asn and p.Leu408Pro) that have only been observed in a single patient. The third group consisted of three variations (p.Lys294Thr(Morimura, et al., 1998), p.Asn321Lys(Thompson, et al., 2000), and p.Ala434Val(Morimura, et al., 1998) that have been suggested to be disease-causing in previous reports (Perrault, et al., 1999; Simovich, et al., 2001; Thompson, et al., 1999), but which receive EPP scores of 0 when using the published algorithm (Stone, 2003).

Table 1 summarizes the calculation of the EPP score for each of the 11 variants in this study. All 11 would be expected to alter an amino acid of the encoded protein, contributing one point to the EPP score. Ten were not observed among a cohort of 200 unaffected Caucasian individuals, contributing a second point to the EPP. Examination of the blosum substitution matrix (Henikoff and Henikoff, 1992) revealed that six of the 11 variations had negative blosum values, which added another point to the EPP for these six variants. A negative blosum value for a specific codon change means that among the hundreds of proteins whose evolutionary variation was studied to create the blosum matrix (Henikoff and Henikoff, 1992), this specific change was less tolerated by evolution than a variation with a non-negative blosum score. We did not have enough data about any of these 11 variants to be able to assert that they were more than 100 times more common in LCA patients than controls and so none of the variants received an EPP point on this basis.

Genotyping of family members of our LCA patients with *RPE65* variations revealed segregation abnormalities for three of the 11 variants in this study. In all three we failed to find a second plausible disease-causing *RPE65* mutation in multiple probands carrying each change. Additionally, for two of these variants (p.Lys294Thr and p.Asn321Lys) we observed a more plausible disease-causing mutation on the same allele in at least one proband. Finally, in one family who carried the p.Ala434Val variant, one of the unaffected parents was found to be homozygous for this change. The observation of these segregation abnormalities caused us to further examine the ethnicity of each of these families and to screen an additional set of control samples based upon the family's self-reported ethnicity. All three variants were found to be present in at least one group of non-Caucasian controls at frequency that is too high for a high-penetrance, disease-causing allele associated with a rare autosomal recessive disease (Table 1). Each of these three variants received an EPP score of zero for this reason.

Figure 1 shows the segregation of putative LCA-causing variations in four of the families studied. The proband (II-1 in Figure 1A) of family A was initially felt to be a compound heterozygote of p.Asn321Lys and p.Arg44Gln. However, when additional family members were examined, we found that the p.Asn321Lys and p.Arg44Gln variations were both present in the proband's father (subject I-1) revealing that the two variations were present on the same allele. Further investigation identified the p.Arg91Trp mutation in the heterozygous state in the proband (II-1) and the unaffected mother (I-2). At this point in the analysis, it was not clear whether either or both of the variations on the paternal allele (p.Arg44Gln and p.Asn321Lys) were pathogenic. We later identified the p.Arg44Gln variant

in the homozygous state in two affected members of a different family (III-1 and III-2 in Figure 1B). The absence of this change in our control subjects made it likely to be disease-causing. In contrast, p.Asn321Lys was found in three of 92 alleles in unaffected Asian Indians (Table 1), which is too common for a high-penetrance, disease-causing allele of a condition that occurs in only 1 in 80,000 people(Stone, 2007).

The p.Ala434Val variation was detected in the heterozygous state in four LCA probands but no second mutation could be identified in these individuals despite careful sequencing of the entire coding sequence of the *RPE65* gene. Moreover, screening the family members (Figure 1C) of one proband revealed p.Ala434Val in the homozygous state in the proband's unaffected mother (I-2). We subsequently identified compound heterozygous mutations in the *CRB1* gene (MIM# 604210) in all three affected members of this family, making *CRB1* the more likely disease gene for this family. When we screened African American control individuals for the p.Ala434Val variant we found it to be present in 4/84 alleles, much too common to be a true LCA-causing variation.

The p.Lys294Thr change was present in the heterozygous state in five LCA probands. In one family (shown in Figure 1D), two additional variations were identified in the *RPE65* gene: p.Ala360Pro and a 22-bp deletion in exon 12 predicted to excise codons 427 to 434 in the mature protein. Molecular analysis of members of this family revealed p.Lys294Thr and the 22bp deletion to lie on the same allele. Later, the p.Lys294Thr change was also detected in 3/52 unaffected Hispanic control alleles (Table 1).

In Vitro Assay of RPE65 Isomerohydrolase Activity

To evaluate independently the pathogenicity of these 11 variants using a biochemical assay, we tested the retinoid isomerase activities of wild-type and mutant RPE65 proteins transiently expressed in 293T-LC cells. Besides RPE65, these cells constitutively express LRAT, which synthesizes all-*trans*-retinyl esters (all-*trans*-REs) from all-*trans*-ROL. These all-*trans*-REs are substrates for the RPE65-isomerase (Mata, et al., 2004; Moiseyev, et al., 2005; Moiseyev, et al., 2003). We determined the isomerase catalytic activities in these cells by measuring the synthesis of 11-*cis*-ROL from all-*trans*-ROL added to the medium. Cells transfected with plasmid expressing wild-type RPE65 synthesized approximately 25 pmols of 11-*cis*-ROL, while cells transfected with non-recombinant pRK5 plasmid synthesized no detectable 11-*cis*-ROL (Figures 2A and 2B). We performed similar isomerase assays using cells transfected with plasmids for the mutant RPE65 proteins. The isomerase activities of RPE65 molecules with p.Gly40Ser, p.Arg44Gln, p.Arg91Gln, p.Thr101Ile, p.Tyr239Asp, p.Arg91Trp, p.Tyr318Asn, and p.Leu408Pro substitutions were all less than six percent of wild-type RPE65 (Table 2). In contrast, RPE65 proteins containing p.Lys294Thr, p.Asn321Lys and p.Ala434Val possessed 68%, 127% and 110% of wild-type Rpe65-isomerase activity, respectively (Table 2), suggesting near-normal isomerase function for these proteins. Results obtained for p.Lys294Thr and p.Ala434Val were similar to those observed previously (Redmond, et al., 2005).

To understand whether the reduced synthesis of 11-*cis*-ROL was due to loss of catalytic activity or decreased expression of the mutant proteins, we measured levels of the wild type and the various mutant RPE65 proteins expressed in 293T-LC cells by quantitative immunoblot analysis. The levels of variants containing p.Arg44Gln, p.Arg91Gln, p.Thr101Ile, and p.Leu408Pro were all at least 70% of wild-type RPE65 (Figure 3A). These results suggest that the reduced production of 11-*cis*-ROL by these variants was due to reduced catalytic activity rather than decreased protein expression. In contrast, levels of RPE65 proteins with p.Gly40Ser, p.Arg91Trp, p.Tyr239Asp and p.Tyr318Asn mutations were significantly reduced compared to wild-type RPE65 (Figure 3A), suggesting that the reduced isomerase activities are due to decreased protein expression. To test if the reduced

levels of mutant RPE65 proteins with p.Gly40Ser, p.Arg91Trp, or p.Tyr318Asn substitutions were due to decreased transfection efficiency or stability of the mRNAs, we performed qRT-PCR on cDNA prepared from transfected 293T-LC cells. Levels of the mutant *RPE65* mRNAs were similar to wild-type (Figure 3B), suggesting that the reduced protein levels are caused by decreased translation efficiency or decreased stability of the mutant RPE65 proteins.

DISCUSSION

The connection between genotype and phenotype is bidirectional. One can study patients affected with a specific disease to discover the causative genes and one can also use the knowledge of discovered genes to help diagnose patients. Although the molecular methods and the nomenclature are very similar in both situations, there are some very substantive differences in the way the information is handled and analyzed depending on the direction of the information flow. For example, the demonstration that a given gene is involved in the pathogenesis of an autosomal recessive inherited eye disease like Leber congenital amaurosis, usually involves the discovery of a number of disease-causing alleles that segregate with the disease phenotype in a number of affected families and that are not present in a control group to any measurable extent. (Morimura, et al., 1998; Thompson, et al., 2000) Many of the disease-causing variations reported in an initial publication of this type are often nonsense mutations, frameshifts, and deletions — that is, variations that seem likely to completely destroy the function of the encoded protein. If it later turns out that one of the putative disease-causing mutations in such a publication is in fact a non-disease causing polymorphism, it does not alter the basic conclusions of the paper in any way because the evidence of the remaining mutations remains sufficiently strong to associate the gene with the disease. In contrast, consider a patient that is contemplating enrollment in a clinical trial of gene replacement therapy or a couple that is contemplating preimplantation genetic testing to lessen the likelihood that their next child will be born blind. In these cases, it would make a very large difference if the genetic variations that had been identified in their DNA were non-disease-causing polymorphisms instead of true disease-causing mutations.

A number of factors can make it difficult to be certain that a genetic variation observed in a given patient is responsible for his or her disease. First, the genome is a noisy place. Two unrelated individuals differ at millions of positions (Wheeler, et al., 2008). Second, many pathogenic mutations are missense variations with seemingly minor effects on the encoded protein. Third, some non-disease-causing polymorphisms are limited to specific ethnic groups. Thus, if the control group is different in ethnic composition from the group of patients being tested, one could mistake a relatively common ethnic-specific polymorphism for a disease-causing mutation. The Leu45Phe polymorphism in RDS (MIM# 179605) is one such ethnic example (Stone, 2003). Fourth, many diseases are caused by mutations in a number of different genes, some of which have not yet been discovered. Thus, in some cases, a true disease-causing mutation observed in the heterozygous state will not play any role in the pathogenesis of a patient's disease.

To help resolve these ambiguities for retinal diseases, we devised an empiric algorithm that predicts the pathogenicity of a given genetic variation in a clinical or laboratory setting. This algorithm results in a score known as the "estimate of pathogenic probability" (EPP) that ranges from zero for non-disease-causing polymorphisms to three for variations very likely to be disease-causing. In the present study, we took advantage of the recent discovery that RPE65 is the isomerase in the visual cycle for regenerating 11-*cis*-RAL chromophore (Jin, et al., 2005; Moiseyev, et al., 2005; Redmond, et al., 2005) to test this algorithm empirically.

The catalytic activity of RPE65 can be assayed in cultured cells by expressing it along with LRAT, adding all-*trans*-ROL to the medium, and measuring production of 11-*cis*-ROL (Jin, et al., 2005). We performed site-directed mutagenesis to generate expression constructs for RPE65 proteins with the same amino-acid substitutions previously identified in a cohort of patients with LCA. The 11 missense variants chosen represent the full range of *RPE65* polymorphisms encountered clinically. Some variants have been seen in several families while others were only seen in single patients. Some alter codons in a way infrequently tolerated by evolution (Arg to Trp and Leu to Pro, both with a blosum value of -3), while others alter codons in a way that is usually well tolerated (Arg to Gln with a blosum value of +1). Importantly, we included three variations (p.Lys294Thr, p.Asn321Lys, and p.Ala434Val) originally thought to be disease-causing (Morimura, et al., 1998; Thompson, et al., 2000) but which subsequently exhibited segregation abnormalities and high allele frequencies in ethnic controls that resulted in EPP scores of zero. In all 11 cases, the assayed isomerase activities were in agreement with the EPP scores. For the eight substitutions predicted by the algorithm to be disease-causing, the enzyme activities were less than six percent of wild type (Table 2). These findings are also in agreement with a previous report showing impaired protein stability of one of the tested variants (p.Arg91Trp) (Takahashi, et al., 2006). In contrast to variants with low predicted pathogenicity based on the EPP algorithm, the substitutions predicted to be non-disease-causing had enzyme activities that ranged from 68-127% of wild-type (Table 2). Interestingly, two substitutions in the latter group had enzyme activities that were significantly higher than wild-type RPE65. In some genetic disorders, a hyperfunctioning allele can be pathogenic. However, in the case of *RPE65*-associated LCA, blindness is caused by a paucity of visual chromophore. Thus, it is unlikely that increased isomerase activity would cause vision loss at birth, the typical phenotype of patients affected with LCA. It is also interesting to note that a 68% reduction in RPE65 enzyme activity appears to be tolerated by the RPE *in vivo*.

The availability of DNA samples from many unaffected individuals of different ethnic backgrounds played an important role in predicting the pathogenicity of the 11 *RPE65* alleles studied here. In the future, very low cost sequencing methods may allow control data to be gathered from several hundred individuals with dozens of ethnic backgrounds. However, until such data are routinely available, geneticists will need to remain alert for three common features of non-disease-causing polymorphisms that were observed in this study: (a) homozygosity in an unaffected individual, (b) the recurrent inability to find a second disease allele in a patient affected with a recessive disease, and (c) the presence of a more plausible disease-causing variation on the same allele.

In summary, we have shown a strong correlation between the results of an *in vitro* assay of enzyme activity and a previously described empirical algorithm for estimating the pathogenicity of observed genetic variants. Although functional assessment of specific disease alleles can be very useful, in many situations it will not be possible to evaluate questionable alleles with this approach. The EPP algorithm provides a practical means of predicting pathogenicity in these settings.

Acknowledgments

REFERENCES

Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, Williams N, Warner J, Campbell H, Porteous ME. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat. 2008; 29(3):367–74. others. [PubMed: 18033691]

Bereta G, Kiser P, Golczak M, Sun W, Heon E, Saperstein D, Palczewski K. Impact of Retinal Disease-Associated RPE65 Mutations on Retinoid Isomerization. Biochemistry. 2008; 47:9856–9865. [PubMed: 18722466]

Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. Hum Mutat. 2008; 29(11):1265–72. [PubMed: 18951437]

Gu S, Thompson D, Srikumari C, Lorenz B, Finckh U, Nicoletti A, Murthy K, Rathmann M, Kumaramanickavel G, Denton M. Mutations in RPE65 cause autosomal recessive childhood-onset severe retinal dystrophy. Nat Genet. 1997; 17:194–197. others. [PubMed: 9326941]

Henikoff S, Henikoff J. Amino acid substitution matrices from protein blocks. Proc Nat Acad Sci USA. 1992; 17:139–141.

Jin M, Li S, Moghrabi W, Sun H, Travis G. RPE65 is the retinoid isomerase in bovine retinal pigment epithelium. Cell. 2005; 122:449–59. [PubMed: 16096063]

Jin M, Yuan Q, Li S, Travis G. Role of LRAT on the retinoid isomerase activity and membrane association of Rpe65. J Biol Chem. 2007; 282:20915–20924. [PubMed: 17504753]

Marlhens F, Bereil C, Griffoin J, Zrenner E, Amalric P, Eliaou C, Liu S, Harris E, Remond T, Arnaud B. Mutations in RPE65 cause Leber's congenital amaurosis. Nat Genet. 1997; 17:194–197. others. [PubMed: 9326941]

Mata N, Moghrabi W, Lee J, Bui T, Radu R, Horwitz J, Travis G. Rpe65 is a retinyl ester binding protein that presents insoluble substrate to the isomerase in retinal pigment epithelial cells. J Biol Chem. 2004; 279:635–643. [PubMed: 14532273]

Mata N, Radu R, Clemmons R, Travis G. Isomerization and oxidation of vitamin a in cone-dominant retinas: a novel pathway for visual-pigment regeneration in daylight. Neuron. 2002; 36:69–80. [PubMed: 12367507]

Moiseyev G, Chen Y, Takahashi Y, Wu B, Ma J. RPE65 is the isomerohydrolase in the retinoid visual cycle. Proc Natl Acad Sci U S A. 2005; 102:12413–12418. [PubMed: 16116091]

Moiseyev G, Crouch R, Goletz P, Oatis JJ, Redmond T, Ma J. Retinyl esters are the substrate for isomerohydrolase. Biochemistry. 2003; 42:2229–2238. [PubMed: 12590612]

Morimura H, Fishman G, Grover S, Fulton A, Berson E, Dryja T. Mutations in the RPE65 gene in patients with autosomal recessive retinitis pigmentosa or leber congenital amaurosis. Proc Nat Acad Sci U S A. 1998; 95:3088–3093.

Perrault I, Rozet JM, Ghazi I, Leowski C, Bonnemaison M, Gerber S, Ducroq D, Cabot A, Souied E, Dufier JL. Different functional outcome of RetGC1 and RPE65 gene mutations in Leber congenital amaurosis. Am J Hum Genet. 1999; 64(4):1225–8. others. [PubMed: 10090910]

Redmond T, Poliakov E, Yu S, Tsai J, Lu Z, Gentleman S. Mutation of key residues of RPE65 abolishes its enzymatic role as isomerohydrolase in the visual cycle. Proc Nat Acad Sci U S A. 2005; 102:13658–13663.

Simovich MJ, Miller B, Ezzeldin H, Kirkland BT, McLeod G, Fulmer C, Nathans J, Jacobson SG, Pittler SJ. Four novel mutations in the RPE65 gene in patients with Leber congenital amaurosis. Hum Mutat. 2001; 18(2):164. [PubMed: 11462243]

Stone E. Finding and interpreting genetic variations that are important to ophthalmologists. Trans Am Ophthalmol Soc. 2003; 101:437–484. [PubMed: 14971589]

Stone E. Leber Congenital Amaurosis-A Model for Efficient Genetic Testing of Heterogeneous Disorders: LXIV Edward Jackson Memorial Lecture. Am J Ophthalmol. 2007; 144:791–811. [PubMed: 17964524]

Takahashi Y, Chen Y, Moiseyev G, Ma J. Two point mutations of RPE65 from patients with retinal dystrophies decrease the stability of RPE65 protein and abolish its isomerohydrolase activity. J Biol Chem. 2006; 281:21820–6. [PubMed: 16754667]

Thompson D, Gyurus P, Fleischer L, Bingham E, McHenry C, Apfelstedt-Sylla E, Zrenner E, Lorenz B, Richards J, Jacobson S. Genetics and phenotypes of RPE65 mutations in inherited retinal degeneration. Invest Ophthalmol Vis Sci. 2000; 41:4293–4299. others. [PubMed: 11095629]

Thompson DA, Lorenz B, Gyürüs P, Gao J, Hanna DB, Andrassi M. RPE65 Mutations In Childhood-Onset Severe Retinal Degenerations. Invest.Ophthalmol.Vis.Sci. 1999; 40:S75.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008; 452(7189):872–6. others. [PubMed: 18421352]

Table 1

Characteristics of 11 selected missense variations observed in LCA patients.

Variant	Frequency in LCA Alleles	Frequency in Caucasian Control Alleles	Blosum Value	Frequency in Ethnic Control Alleles	Segregation Abnormality	EPP
p.Gly40Ser	4/1286	0/400	0			2
p.Arg44Gln	3/1286	0/400	1			2
p.Arg91Trp	11/1286	0/400	-3			3
p.Arg91Gln	3/1286	0/400	1			2
p.Thr101Ile	1/1286	0/400	-1	0/76		3
p.Tyr239Asp	3/1286	0/400	-3			3
p.Lys294Thr	5/1286	0/400	-1	3/52	A, B	0
p.Tyr318Asn	1/1286	0/400	-2	0/94		3
p.Asn321Lys	5/1286	1/400	0	3/92	A, B	0
p.Leu408Pro	2/1286*	0/400	-3	0/64		3
p.Ala434Val	4/1286	0/400	0	4/84	A, C	0

A - unable to identify a second allele in more than one proband.

B - in coupling with a more plausible disease-causing allele.

C - observed homozygously in an unaffected family member.

*
observed homozygously in one LCA proband.

Table 2

Correlation of the estimate of pathogenic probability (EPP) and the *in vitro* isomerohydrolase activity of selected amino acid substitutions in *RPE65*.

Variant	EPP	Activity	
		% of WT	SD (n=3)
WT		100	
pRK5		1.12	0.084
p.Gly40Ser	2	1.65	0.125
p.Arg44Gln	2	1.34	0.161
p.Arg91Trp	3	5.08	0.538
p.Arg91Gln	2	1.33	0.092
p.Thr101Ile	3	1.27	0.076
p.Tyr239Asp	3	4.59	0.414
p.Lys294Thr	0	68	4.52
p.Tyr318Asn	3	5.12	0.747
p.Asn321Lys	0	127	11.1
p.Leu408Pro	3	5.86	0.594
p.Ala434Val	0	110	16.06

Variant - identity of normal or mutated Rpe65 expressed in the 293T-LC cells

WT B normal Rpe65 expressed by pRK5 vector in the 293T-LC cells

pRK5 - empty vector

EPP - estimate of pathogenic probability (see text)

Activity - isomerohydrolase activity of RPE65 with indicated mutation relative to that of WT in an *in vitro* biochemical assay.