Deducing the pathogenic contribution of recessive \textit{ABCA4} alleles in an outbred population

Emily I. Schindler\textsuperscript{1,2}, Erik L. Nylen\textsuperscript{2,3}, Audrey C. Ko\textsuperscript{1,2}, Louisa M. Affatigato\textsuperscript{1,2}, Andrew C. Heggen\textsuperscript{1,2}, Kai Wang\textsuperscript{4}, Val C. Sheffield\textsuperscript{1,3} and Edwin M. Stone\textsuperscript{1,2,*}

\textsuperscript{1}Howard Hughes Medical Institute, \textsuperscript{2}Department of Ophthalmology and Visual Sciences, \textsuperscript{3}Department of Pediatrics and \textsuperscript{4}Department of Biostatistics, University of Iowa, Iowa City, IA, USA

Received June 28, 2010; Revised June 6, 2010; Accepted July 5, 2010

Accurate prediction of the pathogenic effects of specific genotypes is important for the design and execution of clinical trials as well as for meaningful counseling of individual patients. However, for many autosomal recessive diseases, it can be difficult to deduce the relative pathogenic contribution of individual alleles because relatively few affected individuals share the same two disease-causing variations. In this study, we used multiple regression analysis to estimate the pathogenicity of specific alleles of \textit{ABCA4} in patients with retinal phenotypes ranging from Stargardt disease to retinitis pigmentosa. This analysis revealed quantitative allelic effects on two aspects of the visual phenotype, visual acuity ($P < 10^{-3}$) and visual field ($P < 10^{-7}$). Discordance between visual acuity and visual field in individual patients suggests the existence of at least two non-\textit{ABCA4} modifying factors. The findings of this study will facilitate the discovery of factors that modify \textit{ABCA4} disease and will also aid in the optimal selection of subjects for clinical trials of new therapies.

\section*{INTRODUCTION}

A major challenge and opportunity in human genetics is the use of genetic information to accurately predict the natural history (clinical course) of heritable disorders. There are at least five reasons why this is important to do. First, it allows clinicians to give a patient a more accurate prognosis than is possible with clinical methods alone. Second, as new treatments become available for clinical trial, it helps physicians judge the risk/benefit ratio of the treatment for each individual patient. Third, for some diseases, it can help determine the optimal point in the disease to initiate treatment. Fourth, it allows clinical trial investigators to balance treatment and control groups with respect to disease severity—an essential step if one is to obtain an accurate assessment of an intervention’s value with the lowest possible number of subjects and in the shortest possible time. Finally, when one can reliably identify the portion of a given patient’s disease caused by the disease gene itself, it makes it possible to identify additional genes and environmental factors that modify the disease in that patient. Identification of a major modifying factor can serve as the basis for the development of a new treatment.

For dominant diseases, it is often relatively straightforward to determine the relative pathogenicity of individual alleles because each affected person harbors only a single allele and because there are often several affected members of each family available for study. However, for many recessive diseases in outbred populations, the majority of affected patients are compound heterozygotes and the number of affected individuals in each family is small (1–3). As a result, the pathogenic effect of one of the alleles in each genotype is somewhat obscured by the other (4).

Variations in \textit{ABCA4} are the most common cause of autosomal recessive retinal disease in humans. \textit{ABCA4} mutations were originally discovered in patients with autosomal recessive Stargardt disease (SD) but shortly thereafter were also found to cause cone dystrophy, cone-rod dystrophy (CRD) and retinitis pigmentosa (RP), suggesting that a predictable relationship may exist between \textit{ABCA4} genotype and disease severity (5–10). However, the \textit{ABCA4} gene spans more than 100 000 bases and...
several hundred sequence variations have been identified within it (11–14). Most individuals with ABCA4-associated retinal disease are compound heterozygotes (3,4), making it difficult to deduce the specific pathogenic contribution of individual alleles.

In this study, we used multiple regression analysis to deduce the relative pathogenic effect of 16 specific alleles of the ABCA4 gene in patients with molecularly confirmed ABCA4 disease. We also estimated the contribution of non-ABCA4 factors to the disease phenotypes of these individuals.

RESULTS

The primary goal of this study was to determine the relative pathogenicity of different ABCA4 alleles and we therefore designed a screening strategy that would efficiently identify a large number of patients with two disease-causing alleles. A two-step screening method was used to accomplish this. The first step of the screen consisted of an inexpensive, high-throughput allele-specific assay designed to detect 32 of the most common disease-causing ABCA4 variations in the North American population. The second step consisted of automated DNA sequence analysis of the ABCA4 coding sequences of a subset of the individuals found to harbor only a single disease allele during the first step (see Materials and Methods). Since variations in ABCA4 cause a smaller fraction of RP than CRD or SD, the assay was used to screen 457 patients with RP, 60 patients with CRD, 176 patients with SD and 272 normal control individuals. One hundred sixty-nine instances of 61 different disease alleles were observed among the Stargardt patients compared with only four instances of three alleles observed among the control individuals ($P < 10^{-10}$). Similarly, 14 different instances of 10 different alleles were observed among CRD patients ($P < 10^{-6}$ compared with controls) and 43 instances of 23 different alleles were observed among patients with RP ($P < 10^{-5}$).

Eighty individuals with retinal disease and no control individuals were found to have two disease-causing alleles. Parents or other family members were available to confirm the phase of the two alleles in 44 of the 80 unrelated patients with two alleles and inheritance of one disease allele from each parent was confirmed in all of these cases. The specific genotypes of all individuals found to harbor one or more disease-causing variants in ABCA4 are shown in the Supplementary Material, Table.

Table 1. Regression coefficients for visual field score and visual acuity for the most common alleles in the cohort of patients with two disease-causing ABCA4 alleles. Alleles are shown in the order of decreasing visual field score coefficients.

<table>
<thead>
<tr>
<th>Allele</th>
<th>VF model</th>
<th>Acuity model</th>
<th>Occurrences</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu2027Phe</td>
<td>22.81</td>
<td>0.14</td>
<td>4</td>
<td>a</td>
</tr>
<tr>
<td>Leu1201Arg</td>
<td>22.29</td>
<td>0.16</td>
<td>2</td>
<td>a</td>
</tr>
<tr>
<td>Met166fs</td>
<td>20.71</td>
<td>−0.15</td>
<td>4</td>
<td>a</td>
</tr>
<tr>
<td>Gly1961Glu</td>
<td>18.08</td>
<td>0.26</td>
<td>8</td>
<td>a</td>
</tr>
<tr>
<td>Gly863Ala</td>
<td>16.54</td>
<td>0.36</td>
<td>19</td>
<td>a</td>
</tr>
<tr>
<td>Pro1380Leu</td>
<td>15.88</td>
<td>0.39</td>
<td>10</td>
<td>a</td>
</tr>
<tr>
<td>Ala1038Val</td>
<td>15.19</td>
<td>−0.03</td>
<td>12</td>
<td>a</td>
</tr>
<tr>
<td>Leu541Pro</td>
<td>10.95</td>
<td>0.08</td>
<td>1</td>
<td>b</td>
</tr>
<tr>
<td>Asn965Ser</td>
<td>9.3</td>
<td>0.07</td>
<td>3</td>
<td>b</td>
</tr>
<tr>
<td>IVS40 + 5</td>
<td>9.29</td>
<td>0.22</td>
<td>9</td>
<td>b</td>
</tr>
<tr>
<td>Val256Val</td>
<td>9.27</td>
<td>0.84</td>
<td>2</td>
<td>b</td>
</tr>
<tr>
<td>Phe608Ile</td>
<td>7.24</td>
<td>0.48</td>
<td>2</td>
<td>b</td>
</tr>
<tr>
<td>IVS38-10</td>
<td>7.57</td>
<td>0.37</td>
<td>14</td>
<td>b</td>
</tr>
<tr>
<td>Arg1108Cys</td>
<td>1.29</td>
<td>0.81</td>
<td>6</td>
<td>b</td>
</tr>
<tr>
<td>Leu1430fs</td>
<td>0.37</td>
<td>0.6</td>
<td>2</td>
<td>b</td>
</tr>
<tr>
<td>Arg2077Trp</td>
<td>−6.89</td>
<td>0.93</td>
<td>4</td>
<td>b</td>
</tr>
</tbody>
</table>

*When analyzed as groups, A alleles have significantly milder effects on both visual acuity ($P < 10^{-7}$) and visual field ($P < 10^{-7}$) than B alleles (see text).
Figure 1. Correlation of phenotype and genotype. (A) The correlation ($R^2 = 0.56$) of the sum of allelic regression coefficients (genotype pathogenicity score) with the best-corrected visual acuity expressed as the log of the minimum angle of resolution. (B) The correlation ($R^2 = 0.68$) of the sum of allelic regression coefficients with the visual field volume score (see Materials and Methods).

Figure 2. Correlation of visual acuity and visual field. The average visual acuity measurement from each patient is given on the x-axis and the average visual field measurement from each patient is given on the y-axis. The relatively poor correlation between these values ($R^2 = 0.20$) is an indication that non-ABCA4 modifying factors exist that have different effects on these two measures of visual function.
for study was 27 years, so for the patients with multiple visits, the visit that occurred closest to their 27th birthday was the one included in the analyses. To examine the rate of disease progression associated with each genotype, one would ideally have access to multiple visual field and/or acuity measurements from every one of the two-allele patients taken at evenly spaced intervals spanning a period of decades. In the absence of such a complete data set, we attempted to gain a general understanding of the temporal behavior of ABCA4 genotypes by plotting the phenotypic measurements from the cohort of 69 two allele patients as a function of age. Figure 4A shows the visual field data sorted into quintiles of genotype pathogenicity score. Figure 4B shows a similar presentation of the visual acuity data. These figures show that the majority of the phenotypic difference among the quintiles is manifest before age 20. Given that more than 90% of the data in the regression analyses were obtained from patients older than 20 years, it is not surprising that we observed little effect of age on the analyses.

Figure 5 shows the ophthalmoscopic appearance and Goldmann visual fields of five patients from the two allele cohort. In each visual field, the colored lines represent the boundaries (isopters) within which stimuli of various sizes and intensities were visible to the patient. The red line indicates the area in which the smallest and dimmest target (I2e) was visible, whereas the purple line indicates the area within which the largest and brightest target (V4e) was seen (see Materials and Methods). Areas of solid color indicate zones in which a stimulus was not seen. Three individuals (Fig. 5A–F) are concordant with the additive pathogenicity model in that they had small residual values in both the visual acuity and visual field regression analyses (Fig. 3). In contrast, the other two individuals (Fig. 5G–J) each exhibited substantial departures from the additive model in both analyses. The patient shown in Figure 5A and B is typical for individuals with two relatively mild ABCA4 alleles. The visual acuity and visual field are excellent despite numerous yellow flecks at the level of the retinal pigment epithelium (RPE) largely confined to a 3–4 mm ellipse centered on the fovea. Figure 5C and D depicts an individual that is typical for patients with one mild and one severe allele. Flecks are present throughout the posterior pole, a circular patch of RPE atrophy is centered on the fovea and the I2e isopter of the visual field (red line) is moderately constricted. Figure 5E and F depicts an individual with two severe alleles. The fundus shows a ‘beaten metal’ appearance in the macula and extensive bone-spicule-like intraretinal pigment. The I2e isopter has been lost and some areas of absolute scotoma are present. Figure 5G and H depicts a patient with two mild alleles, but visual fields that are considerably worse than average for patients with this genotype pathogenicity score. The RPE is diffusely thin and the choroidal vasculature is easily visible in some areas. The I2e isopter is preserved centrally but is markedly constricted. Figure 5I and J depicts a patient with two mild alleles but visual acuity that is noticeably worse than average for patients with this genotype. The
DISCUSSION

The clinical variation observed among individuals with disease-causing mutations in the same gene can have at least five different non-exclusive explanations: variable structural and/or functional effects of different alleles on the proteins they encode, varying biological effects of the altered proteins on different cell types within affected tissues, effects of one or more modifying genes (i.e. the genetic background), environmental effects and effects of stochastic events occurring during development. For diseases that are inherited in an autosomal recessive fashion, the situation can be even more complicated. When the majority of affected individuals in a population are compound heterozygotes—as is the case when a gene has many alleles and is studied in an outbred population—it is difficult to distinguish the pathological effects of the two alleles in a given patient.

In this study, we investigated a group of patients affected with ABCA4-associated retinal disease, a condition in which most if not all of the previously mentioned complexities exist. Using multiple regression analysis, we found that the majority of the clinical variation among these patients is traceable to the functional sum of their two alleles. The relative pathogenicity of the 16 alleles that were observed most frequently in this cohort is given by the regression coefficients for each allele summarized in Table 1. As additional patients are studied using these methods, these values will become increasingly accurate. However, these numbers, coupled with the demonstration shown in Figure 1 that the alleles behave in an additive fashion with respect to visual phenotypes, will already be useful to investigators who wish to balance treatment and control groups for clinical trials of treatments for ABCA4-associated retinal disease.

We expected that there would be some non-ABCA4-related phenotypic variation among the patients in this cohort and this was most easily seen as the departures from the additive model lines in Figure 1A and B. However, we also expected that much of this variation would affect both visual phenotypes fairly equally. We were surprised to find that the visual acuity and visual field measurements in this cohort had a poorer correlation to each other than either of them did to ABCA4 genotype (Fig. 2), suggesting that the most powerful modifying factors affected the two phenotypes quite differently. This was most dramatically shown by the complete lack of correlation between the residuals of the regression analyses (Fig. 3), which represent the magnitudes of the modifying factors for visual acuity and visual field after subtraction of the effects of the ABCA4 gene itself.

Using the values in Table 1 and Figure 3, we were able to review the clinical records and retinal photographs of the patients in the cohort for clues to the nature of these modifying factors. Three of the patients shown in Figure 5 have small residual values for both the visual acuity and the visual field regression analyses suggesting that their clinical features are quite similar to either the trans or control groups for clinical trials of treatments for ABCA4-associated retinal disease.

The patient with the mildest genotype exhibits noticeable yellowish deposits at the level of the RPE but little evidence of photoreceptor cell death. The patient with the moderate genotype exhibits significant injury of the foveal photoreceptors but insufficient photoreceptor loss elsewhere to result in bone-spicule-like pigmentation. The patient with the most severe genotype exhibits widespread loss of both cone and rod photoreceptors in addition to significant injury to the RPE. This sequence of cellular effects—RPE, cones and rods—is compatible with our current understanding of the function of the ABCA4 protein (recently reviewed by 14) and the physiology of these three cell types. ABCA4 encodes an ATP-binding cassette transporter (15) that couples ATP hydrolysis to the translocation of a visual cycle intermediate (N-retinylidene-phosphatidylethanolamine—N-ret-PE) from the inner leaflet of the photoreceptor disk membrane to the outer leaflet (16–18). Reduced function of this protein allows an additional molecule of all-trans-retinal to covalently and irreversibly bind to...
N-ret-PE to form a more toxic retinal pyridinium compound known as A2E (16,17,19). The outer segments of rod photoreceptors turn over completely every 10 days (20), while cone photoreceptors turn over at about half that rate (21). The outermost parts of the photoreceptor outer segments are phagocytosed by the RPE as new photoreceptive elements are added at the junction of the inner and outer segments. When the accumulation of A2E is fairly slow, as it would be with the mildest genotypes, one might expect its concentration in the photoreceptors to remain below a level that

Figure 5. Fundus photographs and Goldman visual fields of five patients with ABCA4-associated retinal disease. The patients depicted in the upper 10 panels have acuities and visual fields that are completely predicted by the additive effects of their ABCA4 alleles. The patient depicted in (G) and (H) has better visual acuity and poorer visual fields than expected for her genotype, while the patient depicted in (I) and (J) has poorer visual acuity and better fields than expected. (A) and (B) depict the left eye of a 27-year-old woman with a mild ABCA4 genotype, 20/30 acuity and a visual field score of 37.5; (C) and (D) depict the right eye of an 24-year-old man with a moderate ABCA4 genotype, 20/160 acuity and a visual field score of 18.6; (E) and (F) depict the right eye of a 26-year-old man with a severe ABCA4 genotype, 10/160 acuity and a visual field score of 1.1; (G) and (H) depict the right eye of a 47-year-old woman with a moderate ABCA4 genotype, 20/20 acuity and a visual field score of 8.2; (I) and (J) depict the left eye of a 15-year-old woman with a moderate ABCA4 genotype, 20/160 acuity and a visual field score of 43.2.
would cause photoreceptor injury while steadily accumulating in the RPE. With more severe genotypes, one might expect that the first photoreceptor cells to suffer injury would be those with the slowest turnover (cones). Only with the most severe genotypes would the rapidly renewed rods accumulate enough A2E to cause primary (i.e. not secondary to RPE loss) rod photoreceptor death.

The remaining two patients in Figure 5 are illustrative of the phenotypic effects of two unrelated modifying factors: one that modifies acuity with little effect on visual field and one that modifies visual field with little effect on acuity. Patient 4 in Figure 5 has a favorable acuity modifier (protecting the foveal cones in the face of a moderate ABCA4 genotype) and an unfavorable visual field modifier resulting in widespread RPE thinning and patches of RPE atrophy. Patient 5 in Figure 5 is the reverse of patient 4. In this individual, an unfavorable acuity modifier results in early loss of foveal cones, while a favorable visual field protects the RPE cells from death despite significant accumulations of intracellular A2E.

The effect of patient age on the phenotype of ABCA4 disease is interesting. Our data show that the majority of the difference among patients with mild, moderate and severe genotypes is already evident by age 20 (Fig. 4). Although it would be tempting to try to infer the temporal behavior of ABCA4-associated disease based on cross-sectional data, there are two reasons that this might be somewhat misleading. First, there is a definite bias of ascertainment in favor of patients with reduced visual acuity and/or striking fundus findings. Thus, the data points in Figure 4 are unlikely to be a random sampling of all patients affected with ABCA4 retinal disease. Second, two of the patients in this cohort with multiple acuity and field measurements showed substantial worsening after age 40 (data not shown), suggesting that the population trend of stability after age 20 does not apply to all patients.

Application of the methods used in this study to additional cohorts of patients with autosomal recessive disease will further refine our understanding of the relative pathogenic effects of specific alleles. Doing so will improve the prognostic value of genetic testing, contribute to the optimal selection of subjects in clinical trials and facilitate the discovery of non-allelic factors that modify the disease phenotype.

MATERIALS AND METHODS

Subjects

All subjects provided written informed consent for this research study, which was approved by the University of Iowa Human Subjects Committee.
The study included 176 patients with SD, 457 patients with RP, 60 patients with CRD and 272 normal control subjects. Seventeen members of this cohort (13 with SD, 1 with CRD and 3 with RP) had disease-causing mutations identified on one or both alleles in previous studies that employed single-strand conformational polymorphism analysis as the primary screening method (3). All patients and control subjects were ascertained in the outpatient ophthalmology clinic at the University of Iowa. All patients received a complete eye examination including measurement of Snellen visual acuity, slit lamp biomicroscopy of the anterior segment and fundus, and binocular indirect ophthalmoscopy. Most patients had fundus photography and Goldmann perimetry performed as well.

Molecular characterization of the ABCA4 gene

A multi-platform screening approach was used to genotype the ABCA4 gene in all 693 research subjects and 272 controls. Thirty-two of the most common disease-causing ABCA4 variations were selected for the initial screen. The entire research cohort was assayed for these 32 variations using a combination of SNPlex, SSCP analysis, TaqMan and automated DNA sequencing (3). The SNPlex and TaqMan assays were performed as previously described (22,23). Fourteen of the variations were compatible with Applied Biosystems SNPlex allele-specific assay platform. Nine variations were screened by SSCP analysis, three were screened using an Applied Biosystems’ TaqMan assay and six were screened by automated DNA sequencing. All CRD and RP patients who had one plausible disease-causing allele identified in the first tier of screening were assayed by automated sequencing of the entire coding region of the ABCA4 gene in an effort to identify their second disease-causing allele. In addition, Stargardt patients who had one of the three most common alleles (Gly863Ala, Gly1961Glu or IVS38-10) were also sequenced through the entire coding region of the ABCA4 gene.

Visual acuity

The best-corrected visual acuity was recorded in each patient’s medical record as a Snellen fraction (normal = 20/20). Before statistical analysis, these values were converted to the logarithm of the minimal angle of resolution (logMAR) by calculating the base 10 logarithm of the Snellen fraction [e.g. normal = log(20/20) = 0]. For patients with multiple hospital visits, the acuity measurements taken closest to the patient’s 27th birthday were used for the statistical analysis. Values from the left eye and right eye were averaged.

Visual field volume scores

Goldmann visual fields were scanned with a Sharp scanner and analyzed with ImageJ software (available at http://rsbweb.nih.gov/ij/) as follows: transparent layers were added to each field, and the isopters of the visual fields were manually traced onto these layers. Each isopter was assigned a z-axis value according to relative luminous energy of the stimulus (I2e = 100, I3e = 31.7, 14e = 10, III4e = 0.49, V4e = 0.024, no detection = 0). The volume scores were calculated by multiplying the area of each isopter by its associated z-axis value and then summing the values for all isopters. For patients with multiple hospital visits, the visual field measurements taken closest to the patient’s 27th birthday were used for the statistical analysis. Values from the left eye and right eye were averaged.

Statistical analyses

The frequencies of disease-causing alleles observed among the three groups of retinal disease patients were compared with the frequency in controls using Fisher’s exact test. The phenotype of each patient was assumed to result from the additive effect of two alleles and one or more additional factors that were cumulatively represented by a single residual value for each subject. The values for the quantitative contribution of each allele were estimated using a multiple linear regression analysis (24). Specifically, to dissect the effect of each allele, we decomposed the mean phenotype of a person whose genotype consists of allele i and allele j into \( a_i + a_j \), where \( a_i \) is the effect of the i-th allele. To do this, we created a vector \( Y_{\text{vol}} \) of length 51 containing the average logMAR visual acuities for each subject, and a 51x16 matrix, \( X \). Each cell of this matrix, \( X_{ij} \), indicated the number of occurrences of the j-th allele for each subject i. This system of 51 equations was solved with multiple linear regression to give estimates of \( a \), using the statistics program R (available at http://www.R-project.org). To verify the allelic data, each row of \( X \) summed to 2, and each column summed to the corresponding number of occurrences of each allele for this cohort. These calculations were performed in the same manner for \( Y_{\text{vol}} \), the vector of containing each subject’s visual field quantitative phenotype. The estimates of \( a_i \), the coefficients for each allele, for both \( Y_{\text{vol}} \) and \( Y_{\text{vol}} \) are given in Table 1.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

The authors thank Dr Chris Johnson for his helpful suggestions regarding the visual field volume calculations and Jean Andorf and Becky Johnston for their excellent technical assistance.

Conflict of Interest statement. None declared.

FUNDING

This research was supported by the National Eye Institute R01 EY016822, the Foundation Fighting Blindness, the Carver Charitable Trust, the Carver Endowment for Molecular Ophthalmology and the Howard Hughes Medical Institute. Funding to pay the Open Access Charge was provided by the Howard Hughes Medical Institute.
REFERENCES


