Identification of Variants in CNGA3 as Cause for Achromatopsia by Exome Sequencing of a Single Patient

Kevin Lam, MSc; Haiyan Guo, PhD; Graham A. Wilson, MBChB; Susanne Kohl, PhD; Fulton Wong, PhD, MPP

Objective: To report disease-causing mutations in the cyclic nucleotide-gated channel α3 gene (CNGA3) identified by exome sequencing and bioinformatics filtering in a single patient.

Methods: The entire protein-coding sequence of a patient with a retinal disease was enriched by in-solution targeted capture and massively parallel sequenced at 50-fold coverage. The assembled sequence was compared with databases of normal genomic sequences to identify nonsynonymous variants, which were further filtered (1) with a prioritization of genes associated with retinal diseases, (2) according to the likelihood of variant damage to protein function, (3) following the predictions of a recessive model, and (4) against common polymorphisms observed in normal genomes. Clinical evaluation and segregation analysis of the mutant alleles in the patient’s family were performed; mutations were excluded in healthy controls.

Results: Analysis yielded a molecular diagnosis of achromatopsia. Two compound heterozygous mutations were identified in CNGA3 of this patient, c.829C>T p.R277C and c.1580T>G p.L527R; they were not observed in the normal population and cosegregated with the phenotype of achromatopsia in the patient’s family.

Conclusion: These mutations are the cause of achromatopsia in this family.

Clinical Relevance: The key advantages of massively parallel sequencing over linkage mapping and cloning are highlighted by (1) the small sample size required for successful analysis and (2) the rapid and high-throughput manner in which the mutations are identified. This new tool will likely have major effects on the management and research of rare genetic eye diseases in the new era of personalized genomic medicine.

Arch Ophthalmol. 2011;129(9):1212-1217

Over the last quarter century, advances in ophthalmic molecular genetics have been driven to a large extent by new DNA technology. Technological advances in DNA sequencing and widely accessible databases of annotated human genomic sequences are creating a new paradigm for the study of rare ocular genetic diseases. The transforming technology is massively parallel sequencing of the entire genome of individual persons, which is also known as next- or second-generation sequencing.1 While the high cost of this new technology is preventing full realization of its effects at this time, a slight restrictive application is making substantial headway in the analysis of mendelian disorders.2 The new method targets and captures the 180,000 exons (collectively called the exome) dispersed within the human genome for sequencing with the same depth of coverage as whole-genome sequencing and at a much lower cost.3,4 The rationale for this approach is that, while the protein-coding regions compose only 1% of the genome, it is in this portion that 85% of disease-causing mutations occur.5 Thus, massively parallel sequencing of this 1% of human DNA (approximately 30 million bases in length) is a cost-effective and efficient way to identify pathogenic mutations of mendelian diseases. Several recent articles have described the successful application of this method, using for analysis DNA from only a few subjects.3,5-12 To apply whole-exome sequencing and bioinformatics filtering to the study of rare genetic eye disorders, we sequenced the exome of a single patient. As expected, approximately 15,000 genetic variants were identified in this individual’s exome se-
sequence. By applying a few general screening criteria (filters) in the subsequent bioinformatics analysis to separate the background polymorphisms from disease-causing mutations, we aimed to make an accurate molecular diagnosis. The only clinical information used for bioinformatics filtering was that the patient has a retinal disorder. In our study, this general filtering scheme quickly identified 2 heterozygous mutations in the CNGA3 gene, which is known to associate with autosomal recessive achromatopsia.15

METHODS

FRAGMENT LIBRARY CONSTRUCTION

Genomic DNA was isolated from the blood sample using the Gentra Puregene blood kit (Qiagen, Valencia, California). A fragment library of genomic DNA was prepared by using the SOLiD fragment library construction kit (Life Technologies, Carlsbad, California). In brief, 3 µg of genomic DNA was fragmented into lengths ranging from 100 to 150 base pairs (bp) using a Covaris S2 System (Covaris, Woburn, Massachusetts). After end repair, the DNA was ligated with P1 and P2 adaptors from the SureSelect AB adaptor kit (Agilent, Santa Clara, California). Size selection of 150- to 200-bp DNA fragments was performed on a SOLiD library size selection gel (Life Technologies), followed by nick translation and 12-cycle polymerase chain reaction amplification. The quantity and quality of the amplified fragment library was assessed by an Agilent 2100 bioanalyzer before enrichment.

TARGETED CAPTURE AND EXOME SEQUENCING

From the prepared fragment library, 500 ng was enriched for exons by using the SureSelect human all-exon kit (version 1; Agilent). The kit was designed to enrich for all the coding sequences covering a total of 38 megabases in length. In brief, the prepared DNA was in-solution hybridized with SureSelect biotinylated RNA baits for 24 hours at 65°C, and the captured DNA-RNA hybrid library was purified by streptavidin-coated magnetic beads. After digestion of the RNA baits, the captured DNA library was further amplified and used for emulsion polymerase chain reaction according to the manufacturer’s instructions (Life Technologies), based on a library concentration of 0.5pM. The amplified library was then sequenced as single-end 50-bp reads on a SOLiD 3 plus system (Life Technologies). The sample was run on 1 quad of a SOLiD sequencing slide (Life Technologies). A second amplified library was obtained from the same captured library and sequenced independently.

BIOMETRICS ANALYSIS

The 50-bp SOLiD reads were mapped to human genome hg18 in color space using Bowtie version 0.12.3 (http://bowtie-bio.sourceforge.net). Genetic variants were called using SAMtools (http://samtools.sourceforge.net/); nonsynonymous variants were labeled using ANNOVAR (http://www.openbioinformatics.org/annovar/). Further filtering was performed (1) with a prioritization of genes known to associate with retinal disorders (a list of 177 genes from RetNet, http://www.sph.uth.tmc.edu/retinet/); (2) by the PolyPhen-2 Web resource (http://genetics.bwh.harvard.edu/pph2/16 based on high likelihood of damage (pph2 prob >0.9) to protein function caused by a variant, (3) by applying a recessive model, i.e., either 2 heterozygous mutations occurring in the same gene or a homozygous mutation, and (4) by removing common polymorphisms observed in a pool of normal genomes (http://browser.1000genomes.org/index.html).17 Reads generated from the second amplified library were assembled in the same manner as described, and the coverage at the 2 mutation sites was determined.

MUTATION AND SEGREGATION ANALYSES

Exon 7 and flanking intronic sequences of the CNGA3 gene were amplified by polymerase chain reaction with genomic DNA as templates (primer sequences: forward 5'-TCAGAGTGCATTTCGTGATTGT-3' and reverse 5'-GCTTTCAAAGGGTGAGTAGA-3').16 Amplicons were sequenced using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PerkinElmer Applied Biosystems, Foster City, California) and separated on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Sequencing primers were the following: for c.829C>T p.R277C 5'-GCATACGTGTAGCCGAGG-3' and for c.1580T>G p.L527R 5'-GTGCGCAATGCTGGAAGTTCCCC-3'. Editing, sequence alignment, and mutation detection were performed applying the Lasergene Software package (DNASTAR; Lasergene, London, England).

Segregation analysis by DNA sequencing for the presence and independent inheritance of 2 mutant alleles was performed in all family members. Mutations were excluded in 100 healthy (European descent) control samples (200 chromosomes) by direct DNA sequencing.

RESULTS

BIOMETRICS ANALYSIS

From the captured library, 91 346 458 fifty-bp SOLiD reads were generated. Of these, 57 330 833 reads (62.7%) had at least 1 reported alignment to human genome hg18; 56.19% were on target, and 84.30% mapped to only 1 site. The mapped readings provided 49.77-fold coverage (range, 0-5705×) for the targeted exon sequences; 80.81% of bases were covered at more than 7-fold depth (efigure; http://www.archophthalmol.com). Using SAMtools and a cutoff equal to or greater than 20 of the Phred-like quality score, 62 518 variants were called; 15 292 were exonic variants. Of these, ANNOVAR identified 7680 non-synonymous variants; 3533 were not in dbSNP130. The types of variants identified, including indels and splice-site variants, are summarized in eTable 1. Further filtering with a prioritization of genes known to associate with retinal disorders reduced the list to 102 variants associated with 51 genes. Selecting variants on the basis of predicted damage to protein function by PolyPhen-2 resulted in a list of 13 genes and 15 variants. By applying a recessive model, this list was shortened to 5 genes and 7 variants. Screening these 7 variants against the 1000 Genomes Project Consortium database15 identified 3 common polymorphisms previously observed in normal genomes; this fact led to the elimination of a fourth variant based on a recessive model (eTable 2). The list of candidate genes was thus reduced to 2: CNGA3 with 2 heterozygous candidate mutations and TLR3 with 1 homozygous candidate mutation.

Literature review indicated that TLR3 (toll-like receptor 3, also known as CD283) encodes a member of the toll-like receptor family. A recent study by E. Baumann et al18 reported that up-regulation of TLR3 results in increased expression of CNGA3. Therefore, we speculated that TLR3 might be involved in the pathogenesis of this disorder. We found homozygous mutation of TLR3 at c.181G>T p.A61V. This variant (rs1042615) was not observed in 100 control samples, and was not reported in dbSNP130.
toll-like receptor family, which plays a fundamental role in pathogen recognition and activation of innate immunity. It was identified by association study as a candidate gene whose L412F polymorphism might be protective in patients with geographic atrophy related to age-related macular degeneration; on this basis, it was included in RetNet. Therefore, in this context, \textbf{TLR3} is not a retinal disease gene. In contrast, \textbf{CNGA3} is known to associate with achromatopsia, and one of the candidate mutations, c.829C>T p.R277C, has previously been reported. Therefore, knowledge about the 2 candidate genes overwhelmingly supported the hypothesis that the retinal disorder in question is achromatopsia and the proband carries 2 heterozygous mutations: c.829C>T and c.1580T>G, a novel \textbf{CNGA3} mutation (Table, Figure 1, and Figure 2). Sequencing coverage at c.829C>T and c.1580T>G are shown in the Table. All of the reads containing the 2 changes map to only 1 site.

**CASE HISTORY AND CLINICAL FINDINGS**

The proband (II:1; \textbf{Figure 3}) presented shortly after birth with nystagmus. It was noted that she was fascinated by bright lights but performed poorly in a bright environment and appeared to see better in the dark. There was no family history of ophthalmic disease, and the parents were not related. Examination at 9 months of age revealed absence of fixation, and she could not recognize faces in the light. A small-amplitude, rapid, pendular nystagmus was present and there was a paradoxical pupillary response. Findings of anterior segment, ocular media, and fundus examination were normal. The cycloplegic refraction was -8.00 dioptically sphere OU. A clinical diagnosis of achro-

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Reads Covering Mutation, No.</th>
<th>Variant</th>
<th>Reads/ Orientation, No.</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>First amplified library</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.829C&gt;T</td>
<td>38</td>
<td>19 C</td>
<td>5</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19 T</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 T</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>c.1580T&gt;G</td>
<td>34</td>
<td>29 G</td>
<td>21</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>Second amplified library</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.829C&gt;T</td>
<td>28</td>
<td>6 C</td>
<td>3</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22 T</td>
<td>10</td>
<td>+</td>
</tr>
<tr>
<td>c.1580T&gt;G</td>
<td>42</td>
<td>31 T</td>
<td>24</td>
<td>+</td>
</tr>
</tbody>
</table>

©2011 American Medical Association. All rights reserved.

Downloaded From: by a Non-Human Traffic (NHT) User on 11/08/2018
Achromatopsia (rod monochromacy or total color blindness) is a congenital or early-onset retinal disorder with cone photoreceptor function loss. It is estimated to affect 1 in 30,000 individuals worldwide. Clinically it is characterized by reduced visual acuity, pendular nystagmus, photophobia, small central scotomas, eccentric fixation, and reduced or complete loss of color discrimination. In ERG recordings, photopic (cone) responses are absent or markedly diminished, whereas scotopic (rod) responses are essentially normal and remain stable. Genetically, achromatopsia is a heterogeneous condition inherited as an autosomal recessive trait. To date, mutations in 4 genes are known to associate with this disorder.

The primary biological mechanism underlying achromatopsia is a defective cone phototransduction cascade (photopigment → G-protein transducin → phosphodiesterase/cGMP → CNG channel; for review, see the article by Lamb and Pugh). The 4 genes known to associate with achromatopsia are essential components of this cascade, CNGA3 and CNGB3, encoding the channel-forming α and the modulatory β subunit of the cone-specific cyclic nucleotide (cGMP)–gated cation channel (CNG), respectively; GNAT2 is the gene for the cone-specific α subunit of the G-protein transducin.
PDE6C encodes the α′ subunit of the cone-specific phosphodiesterase.\textsuperscript{28,29} In a homologous CNGA3 knockout mouse model, the number of cone photoreceptors is decreased and the remaining cones showed abnormal morphology and degeneration; physiologically, measurable cone function was absent.\textsuperscript{30} Taken together, results obtained from both animal model and human genotype-phenotype correlations strongly support the notion that normal CNGA3 function is critical for cone phototransduction; mutations in the CNGA3 gene that damage CNGA3 protein function lead to achromatopsia.

The mutations identified in this study are most likely causative because of the expected damage to CNGA3 protein function and hence disruption of the cone phototransduction cascade. Arginine 277 is located in the highly conserved S4 transmembrane helix (Figure 2); the R277C mutation has been shown by in vitro expression studies to result in disrupted channel biogenesis and stability.\textsuperscript{31} The R277C mutation is observed recurrently in patients with achromatopsia and their families; it has been suggested that it might be owing to a founder mutation in patients of European descent.\textsuperscript{10} Similarly, the novel mutation L527R affects a highly conserved leucine residue in the cGMP binding domain of the protein (Figure 2), and an apparent clustering of achromatopsia-causing mutations in amino acid positions 510 to 529 has been observed.\textsuperscript{32} Furthermore, another causative mutation that affects L527 and segregates within an unrelated achromatopsia family has been observed recently (S.K.).

In summary, (1) clinical examination of the family, (2) normal scotopic ERG but severely compromised photopic ERG of the proband, (3) mutations deemed to disrupt CNGA3 protein function, and (4) 100% concordance of genotype and phenotype in this family provide strong evidence that the 2 compound heterozygous mutations, p.R277C and p.L527R, are the achromatopsia-causing mutations in this family.

An effective filter used in the present study was limiting the search to genes known to associate with retinal disorders; it streamlined the process to the correct end point. For mutations that occur in genes not already known to be associated with retinal disorders, the same filter would have misguided the process away from the causative mutations. An alternative strategy would be selection of variants that are private to the family, ie, variants observed in the proband’s immediate family but perhaps not in the general population. From this subset, selection of variants that are present only in affected individuals would narrow the search significantly. This strategy was used successfully in identifying mutations underlying Joubert syndrome 2 (mother and affected daughter).\textsuperscript{6}

In addition to comparing private variants of closely related individuals, comparisons can be made of unrelated individuals who have the same clinical phenotype. The rationale is that the same mutation, or different mutations affecting the same gene, may be causative for individuals sharing the same phenotype. This filter was used in the identification of mutations underlying Kabuki syndrome (10 unrelated cases),\textsuperscript{9} Schinzel-Giedion syndrome (4 unrelated cases),\textsuperscript{10} and Fowler syndrome (2 unrelated cases).\textsuperscript{11} The 2 alternative filters described here can be incorporated easily into the high-throughput bioinformatics filtering scheme. Furthermore, the principles outlined here apply equally well to retinal disorders, corneal diseases, or other ocular disease subgroups.

Selecting variants according to the high likelihood of damage to protein function has been used successfully in several studies.\textsuperscript{4} However, there is a graded scale of potential damage; where to draw the line in the high-throughput screening is not always clear-cut. The same difficulty affects filtering by mutation frequency.\textsuperscript{2,6} No doubt, better quantitative models are being developed to address these issues. In the meantime, successful identification of causative mutations for mendelian disorders depends largely on the circumstance of each case; furthermore, different filtering procedures and parameters can lead to the same correct mutation identification.

From the outset, it is clear that exome sequencing would fail to identify some disease-causing mutations (eAppendix). Nevertheless, the odds are in favor of success, especially when viewed in the context of a large number of attempts. When this more restrictive but cost-effective approach fails, the default would be whole-genome sequencing. Several rare diseases have been analyzed successfully by this exhaustive and more expensive method, eg, Miller syndrome,\textsuperscript{32} metachondromatosis,\textsuperscript{33} severe hypercholesterolemia,\textsuperscript{34} and Charcot-Marie-Tooth disease.\textsuperscript{35} Therefore, the new DNA sequencing technology has made it feasible, for the first time, to identify the genetic cause of the rarest forms of disorders, with only single or very few affected individuals.

Submitted for Publication: December 13, 2010; final revision received March 29, 2011; accepted April 4, 2011.

Correspondence: Fulton Wong, PhD, MPP, Departments of Ophthalmology and Neurobiology, Duke University School of Medicine, Durham, NC 27710 (fulton.wong@duke.edu).

Financial Disclosure: None reported.

Funding/Support: This study was supported by Deutsche Forschungsgesellschaft grant KFO134-Ko2176/1-1 (Dr Kohl).

Online-Only Material: The eAppendix, eTables, and eFigure are available at http://www.archophthalmol.com.

Additional Contributions: Britta Baumann provided excellent technical assistance and Huixin Siew participated in bioinformatics analysis.

REFERENCES


