Identification of a GJA3 mutation in a Chinese family with congenital nuclear cataract using exome sequencing

Yi Guo1,2, Lamei Yuan1, Junhui Yi3, Jingjing Xiao2, Hongbo Xu1, Hongwei Lv1, Wei Xiong3, Wen Zheng1, Liping Guan4, Jianguo Zhang3, Hong Xiang1, Yong Qi1 and Hao Deng1,∗

1Center for Experimental Medicine and Department of Neurology, the Third Xiangya Hospital, 2Department of Medical Information, Xiangya School of Medicine, 3Department of Ophthalmology, the Third Xiangya Hospital, 4Cancer Research Institute, Xiangya School of Medicine, Central South University, Changsha, Hunan 410013

Received 29 January 2013; revised 21 May 2013

Congenital cataract, a clinically and genetically heterogeneous lens disorder is defined as any opacity of the lens present at birth. It is responsible for approximately 10% of worldwide childhood poor vision or blindness. To identify the genetic defect responsible for congenital nuclear cataract in a four-generation Chinese Han family, exome and direct Sanger sequencings were conducted and a missense variant c.139G>A (p.D47N) in the gap junction protein-alpha 3 gene (GJA3) was identified. The variant co-segregated with patients of the family and was not observed in unaffected family members or normal controls. The above findings indicated that the variant was a pathogenic mutation. The mutation p.D47N was found in the first extracellular loop (E1) domain of GJA3 protein. Our data suggest that exome sequencing is a powerful tool to discover mutation(s) in cataract, a disorder with high genetic heterogeneity. Our findings may also provide new insights into the cause and diagnosis of congenital nuclear cataract and have implications for genetic counseling.

Keywords: Congenital nuclear cataract, GJA3 gene, Gap junction, p.D47N, Mutation, Exome sequencing

Congenital cataract is defined as any opacity of the lens present at birth. It is responsible for approximately 10% of worldwide childhood poor vision or blindness.1,2 The incidence is estimated to be 2.2-2.49 per 10,000 live births.2 Congenital cataract is a clinically and genetically heterogeneous lens disorder and one third of isolated congenital cataracts are genetically determined.3 Clinically, it can be classified as polar/subcapsular, nuclear, lamellar/zonular, suture, cortical, capsular, and total cataract according to the position of the suffered lens.4,5 Nuclear cataract refers to the opacification limited to the embryonic and/or fetal nuclei of the lens.6,7 Most of them often occur in a non-syndromic autosomal dominant fashion, although autosomal recessive and X-linked inheritance are also known to exist.8

To date, at least 22 disease loci and 17 disease-causing genes for congenital nuclear cataract are reported, including eight crystallin genes (alpha-A-crystallin gene CRYAA, alpha-B-crystallin gene CRYAB, beta-A1-crystallin gene CRYBA1, beta-B1-crystallin gene CRYBB1, beta-B2-crystallin gene CRYBB2, beta-B3-crystallin gene CRYBB3, gamma-C-crystallin gene CRYGC, and gamma-D-crystallin gene CRYGD); one cytoskeletal protein gene (beaded filament structural protein 2 gene BFSP2); three membrane protein genes (gap junction protein-alpha 3 gene GJA3, gap junction protein-alpha 8 gene GJA8 and major intrinsic protein of lens fiber gene MIP); one growth and transcription factor gene (v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog gene MAF); ferritin light chain gene FTL; galactokinase 1 gene GALK1; FYVE and coiled-coil domain containing 1 gene FYCO1; and Nance-Horan syndrome gene NHS.6,9-18 Structural proteins belonging to the gap junction protein family make up the intercellular channels presented in gap junctions and three distinct gap junction protein genes (GJA1, GJA3 and GJA8) are expressed in the lens3.

Earlier, Hansen et al.9 detected crystallin and connexin mutations in 35.71% (10/28) and 21.43%
Danish families, respectively. Here, we report a heterozygous c.139G>A transition (p.D47N) in the GJA3 gene associated with congenital nuclear cataract in a Chinese Han family. The results show that the transition co-segregates with disease phenotype in the family and is absent in controls, indicating it is a pathogenic mutation for congenital nuclear cataract.

Materials and Methods

Clinical evaluation and DNA specimens
A 4-generation, 19-member Chinese Han family from Hunan province, China with autosomal dominant nuclear cataract was recruited from the Third Xiangya Hospital, Central South University (Fig. 1). Thirteen members of the pedigree were involved in this study, including six affected individuals (II:3, II:6, II:8, III:1, III:2 and IV:1) and seven unaffected ones (II:5, II:7, II:9, III:4, III:5, III:6 and III:7). Both affected and unaffected individuals underwent detailed ophthalmic examinations, including visual acuity, slit lamp examination, fundus examination with the dilated pupil and intraocular pressure measurement. The phenotypes were documented by slit lamp photography. One hundred unrelated ethnically-matched subjects (male/female: 50/50, age 40.2 ± 8.3 yrs) without diagnostic features of congenital cataract were recruited from the same region of Mainland China served as normal controls.

Exome capture
Genomic DNA was extracted from peripheral blood using standard phenol-chloroform extraction method. Three micrograms of genomic DNA was used to extract the exome library. DNA of one patient (III:2, Fig. 1) was sheared by sonication and then hybridized to the Nimblegen SeqCap EZ Library for enrichment, according to the instructions from manufacturer. The enriched library targeting the exome was sequenced on the HiSeq 2000 platform to obtain paired-end reads with read length of 90 bp. A mean exome coverage of 78.87× was obtained and such deep coverage provided sufficient depth to accurately call variants at 99.34% of targeted exome.

Read mapping and variant analysis
The human reference genome was obtained from the UCSC database (http://genome.ucsc.edu/), version hg19 (build 37.1). Alignment of the sequences from the patient was performed using SOAPaligner and SNPs were called using SOAPsnp set with the default parameters after the duplicated reads (produced mainly in the PCR step) were deleted. Insertions or deletions (indels) affecting coding sequence or splicing sites were identified. The thresholds for calling SNPs and short indels included the number of unique mapped reads supporting a SNP ≥ 4 and the consensus quality score ≥20. The quality score is a Phred score, generated by the program SOAPsnp 1.05, quality = -10log (error rate). All candidate mutations were filtered against the Single Nucleotide Polymorphism database (dbSNP build 135, http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi), 1000 genomes data (1000genomes release 20100804, http://www.1000genomes.org/), hapmap (2010-08_phase I + II + III, http://hapmap.ncbi.nlm.nih.gov/) and Yan Huang1 (YH1) project.

Sorting intolerant from tolerant (SIFT) prediction (http://sift.jcvi.org/) was performed to evaluate whether an amino acid substitution affects protein function. Sanger sequencing was employed to validate the identified potential disease-causing variants with ABI3500 sequencer (Applied Biosystems, Foster City, CA, USA). Sequences of the primers were as follows: 5'-CGGTGTTTTCATGAGCATTTTC-3' and 5'- GTGG CCCAGGTAGATTGCC-3'.

Results

Clinical findings
The proband (II:3) was 56 yrs old and was diagnosed with bilateral nuclear cataract by slit lamp examination (Fig. 2). According to the medical records, the other five patients (II:6, II:8, III:1, III:2 and IV:1) were diagnosed with bilateral nuclear cataract and had cataract extraction performed. From the hospital...
records, bilateral cataract of all patients was present at birth or developed during infancy, the best corrected visual acuity of the affected eye varied from 0.1 to 0.5 before cataract extraction.

Mutation screening

We sequenced the exome of one patient (III:2, Fig. 1) of the Chinese Han family with congenital nuclear cataract. We generated 7.91 billion bases of sequence from the patient as paired-end 90-bp reads. 7.54 billion bases (95.37%) passed the quality assessment, 6.91 billion bases (91.56%) were aligned to the human reference sequence and 64.55% of the total bases were mapped to the targeted bases with a mean coverage of 78.87-fold

107,890 genetic variants, including 97,408 non-synonymous changes were identified in the patient in the coding regions or the canonical dinucleotide of the splice sites.

A prioritization scheme was applied to identify the pathogenic mutation in the patient, similar to two recent studies23,25. Given that the frequency of congenital nuclear cataract is less than 0.005, we excluded known variants identified in 1000 genome project, HapMap and dbSNP135 with MAF >0.50%. By doing so, we reduced the number of candidate genes by more than 94.32%.

A c.139G>A variant (p.D47N) in GJA3 gene was observed in the patient, while other known disease-causing gene mutations for congenital cataract were excluded (Fig. 3B). The same heterozygous mutation was subsequently identified in all patients with congenital nuclear cataract in the family. The variant co-segregated with patients in the family and was absent in unaffected individuals in this family and in the 100 ethnically-matched unrelated controls.

A multiple amino acid sequence alignment has shown that aspartic acid at position 47 is phylogenetically conserved among different species (Bos taurus, Homo sapiens, Danio rerio, etc.) and gap junctions (human GJA4, GJA8 and GJA10) and SIFT predicts the mutation to be possibly damaging. Our data indicated that p.D47N variant in GJA3 gene was the pathogenic mutation in the family with congenital nuclear cataract.

Discussion

The GJA3 gene is located on 13q11 and consists of two exons encoding a 435-amino acid protein. GJA3 protein is predominantly expressed in the lens, which is essential for maintaining lens transparency1,8. Connexins are a family of structurally-related transmembrane proteins that assemble to form gap junctions and transport metabolites, ions and water in the lens26. All connexins have four transmembrane domains (M1, M2, M3 and M4), two extracellular loops (E1 and E2), a cytoplasmic loop and the NH2-terminus and COOH-terminus located in the intracellular region27.

GJA3 functions as a gap junction that oligomerizes into intercellular channels mediating the transportation of low molecular weight metabolites, ions and second messengers between adjacent cells28. The lens is an avascular structure and lens fiber cells eliminate all intracellular organelles during development to achieve optical clarity. The survival of fiber cells highly depends on intercellular communication29, which is formed mainly by gap junctions. This extensive intercellular communication network is vital for maintaining osmotic and metabolic homeostasis in lens fiber cells and ultimately for transparency of lens30.

In the animal model study, the targeted replacement of GJA8 gene with GJA3 gene in mice has demonstrated that GJA8 gene is required for cell
growth, whereas non-specific restoration of communication by GJA3 gene maintains differentiation. Mutations of GJA3 gene might induce the appearance of a cleaved form of γ-crystallin. The cleavage of γ-crystallin might result in a change in the normal 3-D conformational structure of γ-crystallin, leading to abnormal interactions with other lens proteins. Knock-out of GJA3 gene in different mice strains leads to various degrees of cataracts, similar to the features in this family and other pedigrees, depending on genetic background.

To date, at least 28 GJA3-related pedigrees with congenital cataract, including ours and 23 mutations, including missense, deletion and insertion mutations have been described (Table 1). Clinically, the cataracts share several genotype-phenotype similarities with some inter- and intra-familial differences with respect to the appearance and location of opacities within the lens. Patients from 20 out of 28 families (71.43%) presented with nuclear opacities in the lens. Phenotypes in the other three families were not described. The differences in the GJA3-related cataract phenotypes are reported in different families, even in a same family, which may be explained by the interaction of modifier genes or background environmental factors affecting the expression of the gene and hence resulting in different cataract types.

The GJA3 c.139G>A variant (p.D47N), previously reported in a Chinese congenital nuclear cataract family from Northeast of China was also identified in our congenital nuclear cataract family from South of China, indicating that the D47 might be a mutation hotspot in Chinese population. This variant seemed to be a pathogenic mutation because it segregated with patients with congenital nuclear cataract in our family.

<table>
<thead>
<tr>
<th>Nucleotide change</th>
<th>Amino acid change</th>
<th>Location</th>
<th>Cataract phenotype</th>
<th>Geographic distribution/ethnic background</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.5G&gt;A</td>
<td>p.G2D</td>
<td>NH2-terminus</td>
<td>Nuclear pulverulent and posterior polar</td>
<td>Chinese</td>
<td>32</td>
</tr>
<tr>
<td>c.7G&gt;T</td>
<td>p.D3Y</td>
<td>NH2-terminus</td>
<td>Zonular pulverulent</td>
<td>Hispanic Central American</td>
<td>33</td>
</tr>
<tr>
<td>c.32T&gt;C</td>
<td>p.L11S</td>
<td>NH2-terminus</td>
<td>Ant-egg</td>
<td>Danish</td>
<td>9,34</td>
</tr>
<tr>
<td>c.56C&gt;T</td>
<td>p.T19M</td>
<td>NH2-terminus</td>
<td>Posterior polar</td>
<td>Indian</td>
<td>35</td>
</tr>
<tr>
<td>c.82G&gt;A</td>
<td>p.V28M</td>
<td>M1</td>
<td>Variable</td>
<td>Indian</td>
<td>36</td>
</tr>
<tr>
<td>c.96C&gt;A</td>
<td>p.F32L</td>
<td>M1</td>
<td>Nuclear pulverulent</td>
<td>Chinese</td>
<td>37</td>
</tr>
<tr>
<td>c.98G&gt;T</td>
<td>p.R33L</td>
<td>M1</td>
<td>Finely granular embryonal</td>
<td>Indian</td>
<td>38</td>
</tr>
<tr>
<td>c.130G&gt;A</td>
<td>p.V44M</td>
<td>E1</td>
<td>Nuclear/-</td>
<td>Chinese/Caucasian American</td>
<td>4,30</td>
</tr>
<tr>
<td>c.134G&gt;C</td>
<td>p.W45S</td>
<td>E1</td>
<td>Nuclear</td>
<td>Chinese</td>
<td>39</td>
</tr>
<tr>
<td>c.139G&gt;A</td>
<td>p.D47N</td>
<td>E1</td>
<td>Nuclear</td>
<td>Chinese</td>
<td>3, this study</td>
</tr>
<tr>
<td>c.176C&gt;T</td>
<td>p.P59L</td>
<td>E1</td>
<td>Nuclear punctate/-/-</td>
<td>Caucasian American/Danish</td>
<td>9,40,41</td>
</tr>
<tr>
<td>c.188A&gt;G</td>
<td>p.N63S</td>
<td>E1</td>
<td>Zonular pulverulent</td>
<td>British</td>
<td>42</td>
</tr>
<tr>
<td>c.226C&gt;G</td>
<td>p.R76G</td>
<td>Boundary of E1 and M2</td>
<td>Total</td>
<td>Indian</td>
<td>36</td>
</tr>
<tr>
<td>c.227G&gt;A</td>
<td>p.R76H</td>
<td>Boundary of E1 and M2</td>
<td>Nuclear pulverulent/lamellar and suture</td>
<td>Australian/Danish</td>
<td>9,43</td>
</tr>
<tr>
<td>c.260C&gt;T</td>
<td>p.T87M</td>
<td>M2</td>
<td>Pearl box</td>
<td>Indian</td>
<td>44</td>
</tr>
<tr>
<td>c.559C&gt;T</td>
<td>p.P187S</td>
<td>E2</td>
<td>Nuclear pulverulent</td>
<td>Chinese</td>
<td>46</td>
</tr>
<tr>
<td>c.560C&gt;T</td>
<td>p.P187L</td>
<td>E2</td>
<td>Zonular pulverulent</td>
<td>Caucasian</td>
<td>47</td>
</tr>
<tr>
<td>c.563A&gt;C</td>
<td>p.N188T</td>
<td>E2</td>
<td>Nuclear pulverulent</td>
<td>Chinese</td>
<td>8</td>
</tr>
<tr>
<td>c.616T&gt;A</td>
<td>p.F206I</td>
<td>M4</td>
<td>Nuclear</td>
<td>Chinese</td>
<td>1</td>
</tr>
<tr>
<td>c.1137insC</td>
<td>p.S380fs</td>
<td>COOH-terminus</td>
<td>Zonular pulverulent</td>
<td>British</td>
<td>42</td>
</tr>
<tr>
<td>c.1143_1165del23</td>
<td>p.S381fs</td>
<td>COOH-terminus</td>
<td>Punctate nuclear</td>
<td>Chinese</td>
<td>41</td>
</tr>
</tbody>
</table>

M1, first transmembrane domain; E1, first extracellular loop; M2, secondary transmembrane domain; E2, secondary extracellular loop; M4, fourth transmembrane domain.
and was absent in both unaffected family members and the 100 unrelated ethnically-matched controls.

The mutation p.D47N in the E1 domain of GJA3 may influence the hemi-channel docking and subsequently may lead to damaging interference with conformation and function of GJA3\(^{3,30}\). Except p.D47N, four other reported mutations (p.V44M, p.W45S, p.P59L and p.N63S) located in the E1 domain of GJA3 have been reported in seven pedigrees with congenital cataract\(^{3,30}\). The high frequency mutations (32.14%, 9/28) in the E1 domain of GJA3 protein by meta-analysis of GJA3-related congenital cataract suggested that this domain was the hot mutation region.

In summary, our study shows that exome sequencing is a powerful tool to discover mutation(s) in cataract, a disorder with high genetic and clinical heterogeneity. Our findings may also provide new insights into the cause and diagnosis of congenital cataract and may have implications for genetic counseling in the future.

**Acknowledgements**

We thank the participating patients and investigators for their cooperation and efforts in collecting the genetic information and DNA specimens. This work was supported by grants from National Natural Science Foundation of China (81271921; 81101339; Sheng Hua Scholars Program of Central South University, China (H.D.); Research Fund for the Doctoral Program of Higher Education of China (20110162110026); Science and Technology International Cooperation Key Project of Hunan Province, China (2011WK2011); Construction Fund for Key Subjects of the Third Xiangya Hospital, Central South University, China; Students Innovative Pilot Scheme of Central South University (CL11280); China; Graduate Independent Exploratory Innovative Project of Central South University, China (2013zzts101); the National High Technology Research and Development Program of China - 863 Program (NO.2012AA02A201); the Guangdong Enterprise Key Laboratory of Human Disease Genomics; National Gene Bank Project of China; Shenzhen Key Laboratory of Transomics Biotechnologies (NO.CXB201108250096A) and Shenzhen Engineering Laboratory for Clinical Molecular Diagnostic.

**References**
