Genome-Wide Detection of Copy Number Variations in Unsolved Inherited Retinal Disease

Xiu-Feng Huang, Jian-Yang Mao, Zhi-Qin Huang, Feng-Qin Rao, Fei-Fei Cheng, Fen-Fen Li, Qing-Feng Wang, and Zi-Bing Jin

Division of Ophthalmic Genetics, Lab for Stem Cell & Retinal Regeneration, Institute of Stem Cell Research, The Eye Hospital, Wenzhou Medical University, Wenzhou, China

Correspondence: Zi-Bing Jin, Division of Ophthalmic Genetics, Lab for Stem Cell & Retinal Regeneration, The Eye Hospital, Wenzhou Medical University, Wenzhou 325027, China; jinzb@mail.eye.ac.cn.

Submitted: September 7, 2016
Accepted: December 14, 2016

Citation: Huang X-F, Mao J-Y, Huang Z-Q, et al. Genome-wide detection of copy number variations in unsolved inherited retinal disease. *Invest Ophthalmol Vis Sci*. 2017;58:424–429. DOI:10.1167/iovs.16-20705

Purpose. Inherited retinal diseases (IRDs) are a clinically and genetically heterogeneous group of Mendelian disorders that plays a crucial role in the etiology of blindness across the world. Molecular genetic diagnosis of IRD remains extremely complex and challenging because mutations are only detected in 40% to 60% of cases. In this study, we aimed to dissect the contributions of copy number variations (CNVs) in IRD patients.

Methods. A total of 50 patients were diagnosed with IRD, all of whom previously tested negative for pathogenic mutations in known disease genes. Single-nucleotide polymorphism array analysis was performed by using the HumanCoreExome BeadChip. Analyses of CNVs were carried out by using GenomeStudio, KaryoStudio, and cnvPartition. The putative pathogenic CNVs were further confirmed by real-time quantitative PCR.

Results. We identified four novel CNVs in three different genes (one duplication in *USH2A* gene, two duplications in *CEP290* gene, and one duplication in *RIMS2* gene) in total four families, at a detection rate of 8% (4/50). All of these CNVs are currently absent in all databases. Three variations are located in genes that are already known to cause inherited retinal disease: *USH2A* and *CEP290*, while the association between mutation in the *RIMS2* gene and IRD is reported for the first time.

Conclusions. We performed whole-genome–wide CNV analyses in a large cohort as an alternative approach to molecular diagnosis of IRDs. This study dissected the contributions of CNVs of IRDs, not only increasing the yield in genetic testing but also suggesting the CNVs should be analyzed in the patients with IRDs.

Keywords: inherited retinal diseases, copy number variations, whole-genome wide SNP genotyping, genetic testing, novel genotype-phenotype correlations

Inherited retinal diseases (IRDs) are a clinically and genetically heterogeneous group of Mendelian disorders that plays a crucial role in the etiology of blindness across the world. Inherited retinal diseases are characterized by progressive pathologic change in rod and/or cone photoreceptor cells, which are the components of the ocular responsible for absorbing and converting light into electrical signals.[1,2] Diverse inheritance patterns have been described in families with IRDs, including autosomal recessive, autosomal dominant, X-linked, mitochondrial, and digenic traits.[3] Adding to the complexity of these diseases, the clinical manifestation and course of IRDs vary widely in patients. In fact, mutations in the same gene can lead to either syndromic or nonsyndromic retinal degenerations.[4]

Understanding the genetic basis of IRD patients provides a number of benefits: (1) supplying an accurate prognosis of the clinical course; (2) opportunity for genetic counseling; and (3) potential inclusion in clinical trials of stem cell or gene therapy. Tremendous efforts in genetic diagnostic testing of IRDs have been made, including microarray-based genotyping, targeted exome sequencing, and whole exome sequencing. To date, more than 250 disease-causing genes and 4000 different mutations have been identified in multiple phenotypes of IRDs (RetNet; https://sph.uth.edu/retnet/home.htm, in the public domain).[5,6] Among these identified IRD genes, the top leading ones include *USH2A*, *EYS*, *ABCA4*, *PDE6B*, *RPGR*, and *RHO*.[7–13] However, a definitive molecular diagnosis of IRD remains extremely complex and challenging because mutations are only detected in 40% to 60% of cases.[4,8,14–16] Most of the previous studies of IRDs have focused on single-base substitution mutations or small insertions/deletions; however, it is reasonable to speculate that large duplications or deletions can also contribute to IRDs.

Copy number variations (CNVs) have become increasingly recognized as a potential key genetic cause of Mendelian diseases. Several studies[9,17–20] have since revealed some of the pathogenic CNVs in patients with IRDs. Nevertheless, the contribution and variation spectrum of CNVs in IRD cohort remain unclear. In addition, few studies have attempted to screen for large-scale genome-wide CNVs in both IRD-associated genes and unknown genes.

In this study, we set out to investigate potential CNVs in IRD patients with no detected mutation in known IRD-associated genes. We used whole-genome–wide single-nucleotide polymorphism (SNP) genotyping arrays to analyze CNVs in a cohort of 50 unrelated IRD patients, who have previously tested negative for pathogenic mutations in known IRD genes according to results from targeted exome sequencing.[15] We

TABLE. Identification of Copy Number Variations in Patients With Inherited Retinal Disease

Patient ID	Family	Clinical Diagnosis	Type	Chr	Estimated CNV Coordinate	Size	Copy (Value)	Genes	Related Disease
F1:II:2	F1	RP	S	Chr1	215868952-216173866	304914	Gain (3)	*USH2A*	USH, RP
F2:II:1	F2	LCA	S	Chr12	88444154-88561356	117202	Gain (3)	*CEP290*	LCA
F3:II:1	F3	LCA	S	Chr12	88444154-88561356	117202	Gain (3)	*CEP290*	LCA
F4:II:1	F4	RP	AD	Chr8	104901578-105178819	277241	Gain (3)	*RIMS2*	None

AD, autosomal dominant; Chr, chromosome; S, sporadic; USH, Usher syndrome.

identified potential disease-causing CNVs in both IRD genes and novel genes. Our results indicated the potential role of CNV screening as diagnostics for IRDs and provided insights into the genetic complexity of this group of diseases.

METHODS

Study Subjects

A total of 50 unrelated individuals with IRDs were recruited at the Eye Hospital of Wenzhou Medical University. All subjects underwent standardized ophthalmic examinations. Patients' initial symptoms and complaints included defective vision, night blindness, and narrowed visual field. Informed consent was obtained. Study protocol was conducted in accordance with the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the Eye Hospital of Wenzhou Medical University.

Single-Nucleotide Polymorphism Array Platform

Genomic DNA was extracted from peripheral blood samples from patients and family members by using a Simgen Blood DNA Mini Kit (Simgen, Hangzhou, China) according to the manufacturer's instructions. Total DNA concentrations were determined by using NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). Genomic DNA samples were subjected to SNP array analysis with the HumanCoreExome-24 BeadChip (Illumina, San Diego, CA, USA) and HumanCoreExome-12 BeadChip (Illumina). The arrays contain 547,644 markers and 265,919 exome-focused markers. DNA samples were tested according to manufacturer's instructions including the following steps: DNA digestion, ligation, PCR amplification, fragmentation, labeling, and hybridization. An iScan Reader (Illumina) was used for scanning the array slide. Single-nucleotide polymorphisms were excluded if their call rate was less than 95%.

Analyses and Identifications of CNVs

Analyses of CNVs were performed by using GenomeStudio Software v2011, KaryoStudio Software v1.4, and CNVpartition CNV Analysis Plug-in (v3.2.0; Illumina). The detailed analysis parameters have been described in Supplementary Table S1. Samples with low-data quality would be removed during the quality control (QC) process. The minimum probe count required to call a CNV is 3 (Supplementary Table S1), according to the User Guide of GenomeStudio by Illumina, Inc. The data were analyzed by GenomeStudio and KaryoStudio. However, for the purpose of reducing the possibility of false-positive signals, we used only the ones overlapping. Those with Log R ratio standard deviation > 0.3 were removed from subsequent analyses. The default value of confidence threshold was >0.75. Candidates of gain/loss of copy number regions were filtered by the following public databases: Database of Genomic Variants (DGV; http://dgv.tcag.ca/dgv/app/home, in the public domain). To investigate the inheritance pattern and gene function, Online Mendelian Inheritance in Man (OMIM; http://www.omim.org/, in the public domain) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed, in the public domain) were used.

Validations of CNVs

The putative pathogenic CNVs were further confirmed by real-time quantitative PCR using 7500 Real-Time PCR system (ABI, Carlsbad, CA, USA). The RNaseP was used as an endogenous control. The 20-μL reaction contained 10 μL 2X faststart universal SYBR Green Master (ROX; Roche, Basel, Switzerland), 10 μM of each primer, and 1 μL genomic DNA as template. The following thermal conditions were performed: 10 minutes of preheating at 95°C followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. For normalization, the threshold cycle (CT) values of the endogenous control were subtracted from the corresponding CT values of candidate genes to generate ΔCT values. All relative expression values were reported as means ± standard errors of the means on a 2-log scale.

RESULTS

The CNV analysis pipeline was described in Supplementary Figure S1. After comprehensive analyses of CNVs by a combination of GenomeStudio, KaryoStudio, cnvPartition, and real-time quantitative PCR validations, we identified four novel CNVs in three different genes (one duplication in *USH2A* gene, two duplications in *CEP290* genes, and one duplication in *RIMS2* gene) in four patients at a detection rate of 8% (4/50). The Table summarized the results of CNV analysis in this study (Table). All of these CNVs are currently absent in all available databases. Three of the CNVs are located in genes already known to cause IRDs: *USH2A* and *CEP290*. However, the potential association between mutations in *RIMS2* gene and IRDs has not yet been elucidated.

The first variant, a duplication in *USH2A* gene, was detected in patient F1:II:2, who was diagnosed as having sporadic retinitis pigmentosa (RP; Fig. 1A). Results from Log R ratio and B Allele Freq indicated that an additional gained copy of ~300 kb, spanning exons 33 to 61, was present in *USH2A* gene (Chr1: 215868952-216173866) (Fig. 1B). We used real-time quantitative PCR to verify this reported CNV. The RT-PCR primers were designed to recognize the normal copy regions (exon 2), the CNV regions (exon 48 and exon 59). The real-time quantitative PCR data showed that the gene dosage in patient F1:II:2 was ~1.5 times higher than in control sample regarding exon 48 and 59, while the dosage for exon 2 remained normal (Fig. 2). The results confirmed that patient F1:II:2 has gained an additional copy of the *USH2A* gene, specifically at the region spanning exons 33 to 61. We then referred back to data from targeted exome sequencing and found that patient F1:II:2 also harbored a heterozygous missense mutation in *USH2A* (c.12575G>A, p.R4192H) (Fig. 1C). This is a very rare mutation with frequency of 0.0005

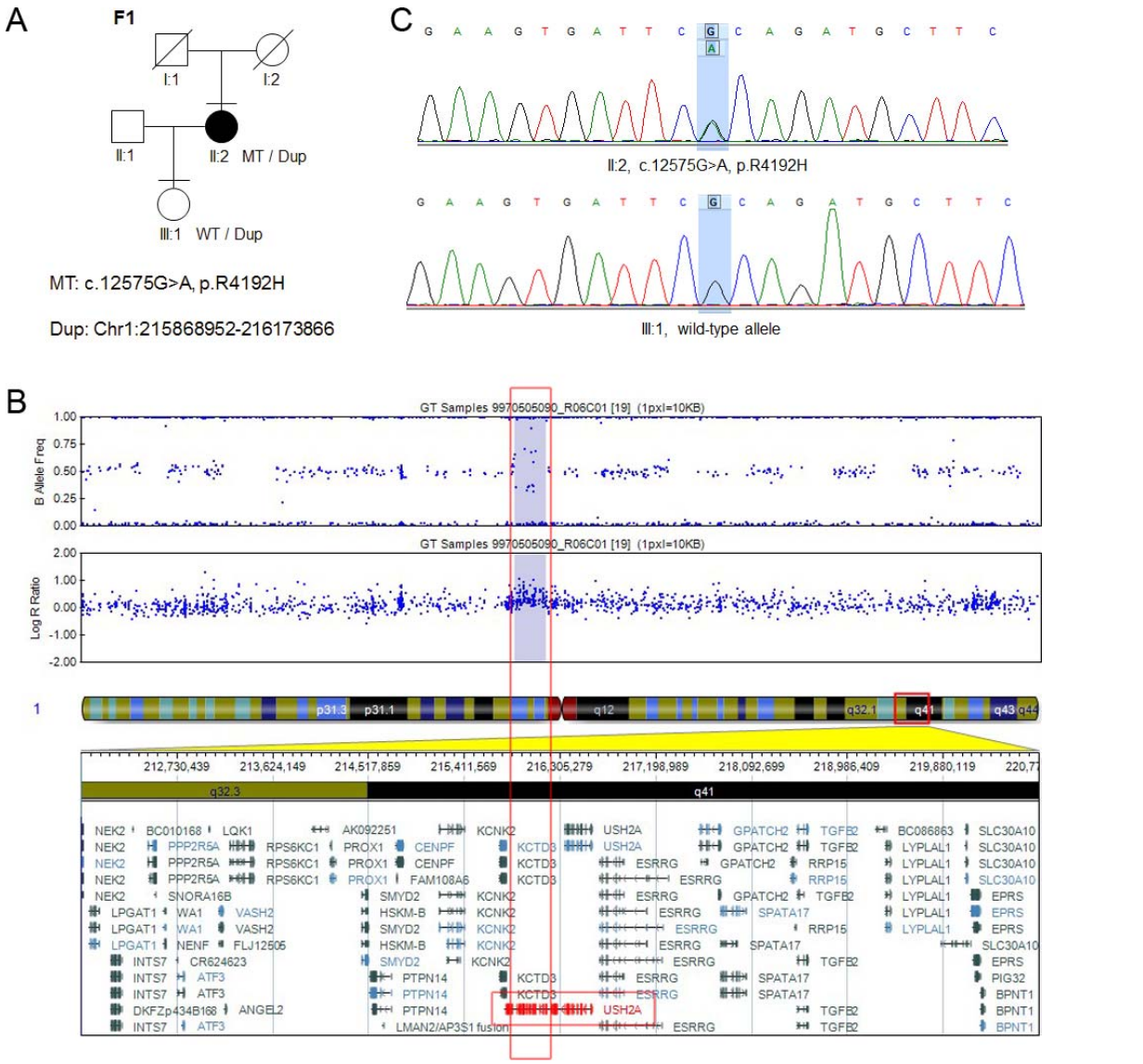

Figure 1. Identification of copy number variation and missense mutation in *USH2A*. (**A**) The pedigree of family F1. (**B**) Single-nucleotide polymorphism array and CNV analysis showed a ~300-kb duplication in *USH2A* gene in F1:II:2. (**C**) The sequence electropherograms of R4192H mutation in *USH2A* gene.

(according to ExAC database), and no homozygote has been identified in a total of 120,484 alleles. Assessment of protein functional effects by MutationTaster indicated this mutation to be deleterious; however, SIFT and Polyhen2 predicted it as tolerable. To validate whether this missense mutation and the CNV were located in two different alleles, the patient's healthy daughter (F1:III:1) was recruited into the study and tested for these mutations. Interestingly, real-time quantitative PCR results confirmed the presence of CNV, and Sanger sequencing revealed a wild-type allele at the R4192 position. These results further confirmed that the CNV and the missense mutation were not present in the same allele, therefore indicating an autosomal recessive inheritance model. Taken together, we identified two compound mutations, a novel CNV and a novel missense mutation, in a sporadic case of RP.

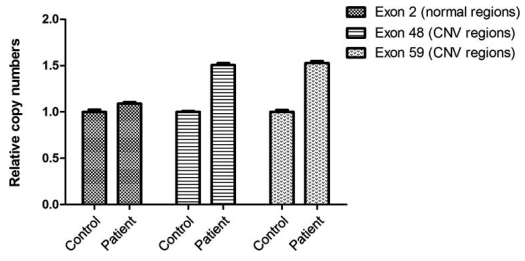

Figure 2. The duplication in *USH2A* gene was confirmed by real-time quantitative PCR.

The second and third variants were identified in two unrelated individuals as the same CNV. This CNV contained an additional ~100-kb gained copy of the *CEP290* gene that spanned nearly the entire gene (Chr12:88444154-88561356) (Figs. 3, 4). Patients F2:II:1 and F3:II:1 were clinically diagnosed as exhibiting sporadic Leber congenital amaurosis (LCA). Notably, both patients carried a heterozygous missense variant in the *CEP290* gene: c.829G>C (p.E277Q) in patient F2:II:1 and c.1991A>G (p.D664G) in patient F3:II:1. Homozygotes of both of these variants were identified in ExAC databases. There are 2 E277Q and 13 D664G homozygotes, indicating that these two variants are polymorphic. Based on

Figure 3. Identification of copy number variation in *CEP290* in patient F2:II:1.

these results, the pathogenicity of this CNV in the *CEP290* gene is still inconclusive.

The fourth variant is a ~277-kb gained copy of the *RIMS2* gene (Chr8:104901578-105178819) in patient F4:II:1, who is from an autosomal dominant RP family (Figs. 5A, 5B). However, as we could not assess the genomic DNA of the affected father, we were unable to determine whether this CNV was inherited from a paternal allele. Moreover, unlike the known causative genes of IRDs such as *USH2A* and *CEP290*, this variant involving *RIMS2* has not been reported in any human disease. Therefore, the role of CNV in the *RIMS2* gene in pathogenesis of IRDs remains unclear.

Discussion

Human Mendelian diseases can be caused by different types of genetic defects, including single-base substitution mutations (nonsense, missense, alternative splicing, and intronic mutations), small insertions/deletions, and CNVs. Of note, single-base substitution mutations and small insertions/deletions could be efficiently identified by using linkage analysis and diverse sequencing techniques, whereas the identification of CNVs requires different methodologies. Although thousands of mutations have been revealed in patients with IRDs, very few of them are CNVs. As the current detection rate of molecular genetic diagnosis in IRDs is only at 40% to 60%, it is logical to presume that a portion of IRD-associated mutations is being overlooked, and that a large number of CNVs could potentially be identified in cases of IRDs.

Here, we used genome-wide SNP genotyping techniques to investigate the CNVs in an IRD cohort. We have previously identified the genetic causes of 99 cases in a total of 179 families with IRDs by using targeted exome sequencing.[15] Among those unsolved IRD families, 50 probands were recruited into this study. We identified three different CNVs in four patients (detection rate: 8%, 4/50). All of the reported CNVs were absent in DGV and any of the other databases. Patient F1:II:2 harbored special compound mutations: a ~300-kb CNV and a missense mutation in *USH2A*. From the genotyping results of her family members, we demonstrated that these two variants were located in different alleles, and therefore, the disease was inherited in an autosomal recessive fashion. Our results also indicated the deleterious impact of

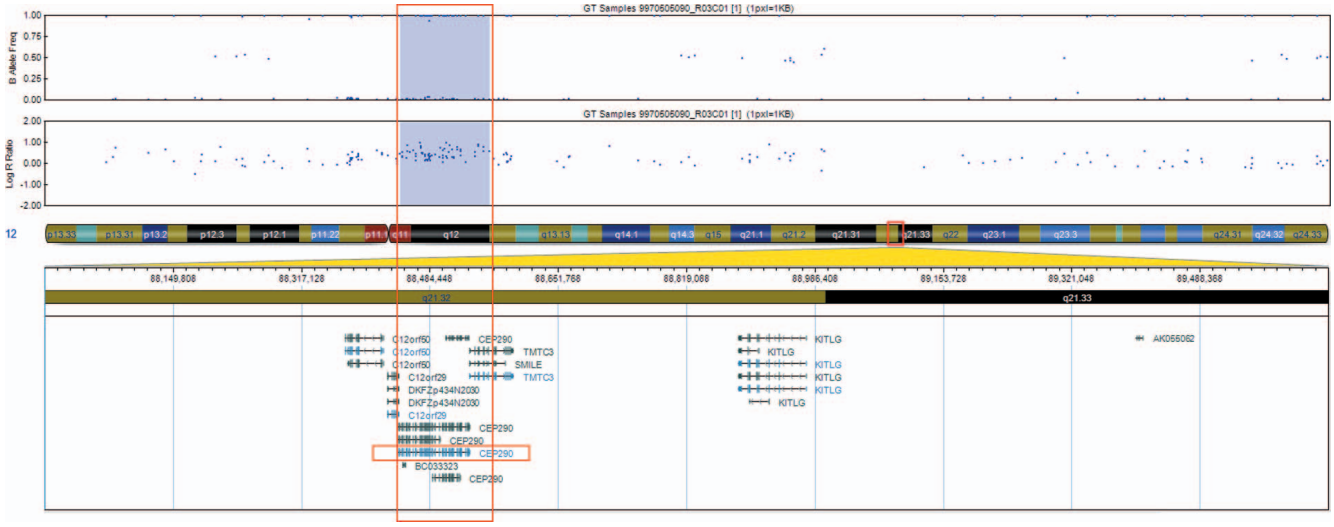

Figure 4. Identification of copy number variation in *CEP290* in patient F3:II:1.

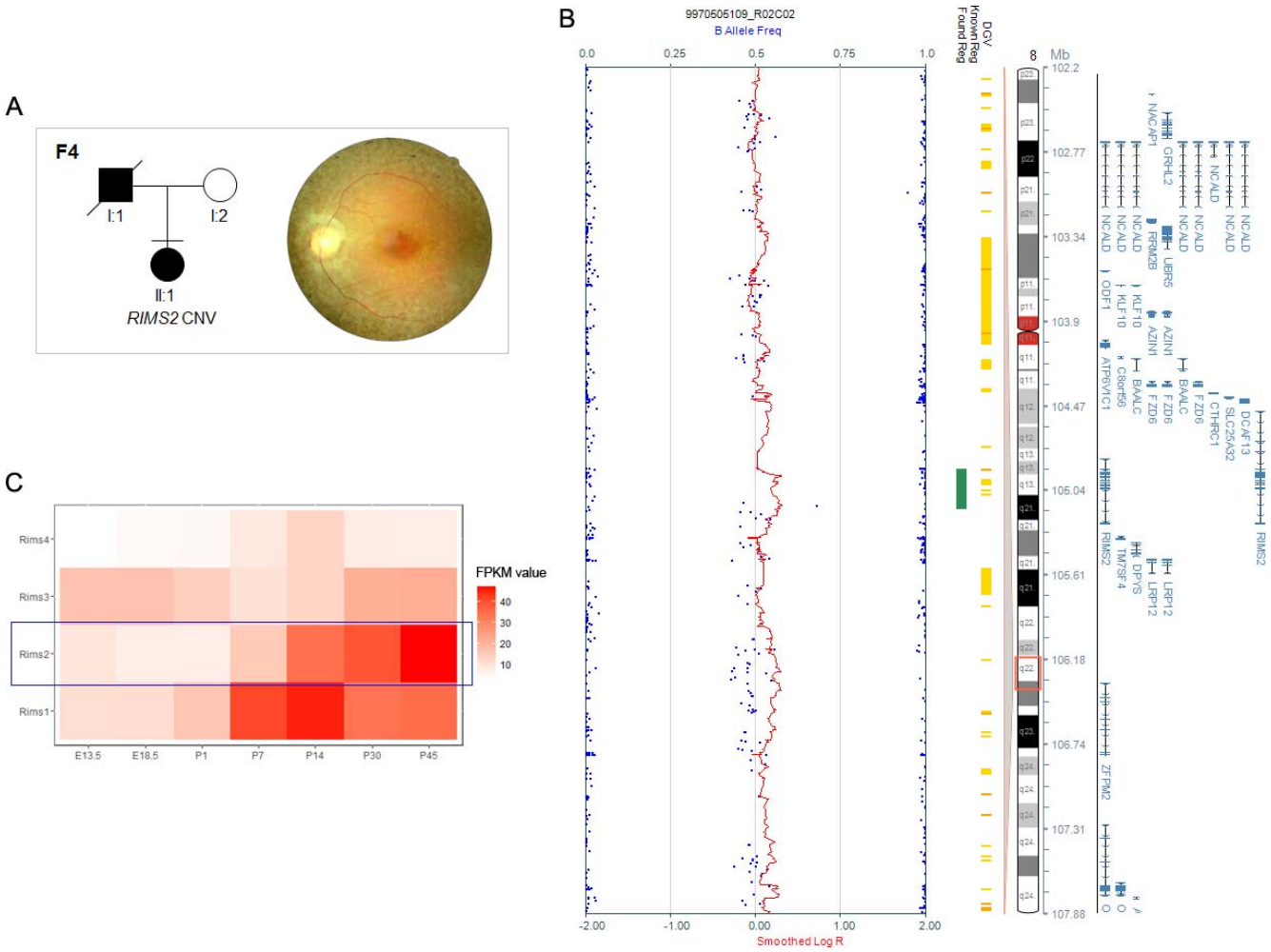

Figure 5. Identification of copy number variation in *RIMS2*. (**A**) The pedigree of family F4 and the fundus of patient F4:II:1. (**B**) Single-nucleotide polymorphism array and CNV analysis showed a ~277-kb duplication in *RIMS2* gene in F4:II:1. (**C**) *Rims2* is expressed as early as stage E13.5 and expression increases during postnatal development of the mouse retina.

this large duplication. Notably, mutations in *USH2A* could lead to Usher syndrome or RP.[21] In this study, patient F1:II:2 was clinically diagnosed to suffer from RP. Our data also provide novel genotype–phenotype correlations in cases of IRDs.

Unfortunately, the functional impact of the other three CNVs remains unclear. In families 2 and 3 (F2 and F3), both probands carried an identical gained copy (~100 kb) and a missense mutation unique to each patient (E277Q in patient F2:II:1; D664G in patient F3:II:1) in *CEP290*. As *CEP290* is the major disease-causing gene of LCA, results of genotype analysis were consistent with the patients' clinical features.[22] However, both of the missense mutations are present as homozygotes in exome database (according to ExAC), suggesting that these mutations do not affect protein function. Therefore, the pathogenicity of this CNV in *CEP290* remains to be investigated in future study.

In the fourth case, we discovered a novel duplication in *RIMS2* gene, which has yet to be associated with retinal development. *RIMS2* is a member of the RAS gene superfamily that regulates synaptic vesicle exocytosis.[23] Mutations in *RIMS1*, which also belongs to this gene family, cause autosomal dominant cone–rod dystrophy.[24] Interestingly, disease in this particular patient also followed an autosomal dominant mode of inheritance. Notably, according to our in-house database of mouse retina RNA-Seq, *Rims2* is expressed as early as stage E13.5 and expression increases during postnatal development of the mouse retina (Fig. 5C). However, as the genomic DNA of the affected father was unavailable, we were unable to perform cosegregation testing. To further elucidate the functional impact of this CNV in *RIMS2* gene would require an expanded screening of an independent patients' population.

Recent studies have been performed to elucidate the role of CNVs in patients with IRDs.[9,17–20,25] Compared to those previous studies, our study possessed two advantages: (1) we were the first to use whole-genome–wide genotyping arrays to dissect the contributions of CNVs in a large cohort of IRD patients; and (2) we analyzed not only the deletions, but also the duplications of genetic materials. However, there were certain limitations in this study as well. Firstly, the sample size was small, which could explain the low number of positive results. Secondly, with the exception of *USH2A*, the roles of the reported variants in IRDs are still unknown. Independent replication and functional experiments are necessary to further determine the contribution of CNVs in IRD pathogenesis.

In summary, we performed a whole-genome–wide search of CNVs by using genotyping arrays in a large cohort of IRD patients. We identified four novel CNVs in three different genes, including a novel candidate gene for IRDs. In addition, novel genotype–phenotype correlations of IRDs were described. Finally, not only does this study expand the potential of genetic diagnosis of IRD by examining the contributions of

CNVs, it also suggests that CNVs should be analyzed in patients with unsolved IRDs.

Acknowledgments

The authors thank the families for participation in this study.

Supported by the National Key Basic Research Program (2013CB967502 to Z-BJ), National Natural Science Foundation of China (81371059, 81522014 to Z-BJ), Zhejiang Provincial Natural Science Foundation of China (LR13H120001 to Z-BJ), NHFPC Grant-in-Aid for Medical and Health Science (201472911 to Z-BJ), Wenzhou Science and Technology Innovation Team Project (C20150004), Innovation Research Program of the Eye Hospital (YNCX201511), and Research Program of Zhejiang Provincial Department of Education (Y201534214).

Disclosure: **X.-F. Huang**, None; **J.-Y. Mao**, None; **Z.-Q. Huang**, None; **F.-Q. Rao**, None; **F.-F. Cheng**, None; **F.-F. Li**, None; **Q.-F. Wang**, None; **Z.-B. Jin**, None

References

1. Wright AF, Chakarova CF, Abd El-Aziz MM, Bhattacharya SS. Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. *Nat Rev Genet*. 2010;11:273–284.
2. den Hollander AI, Black A, Bennett J, Cremers FP. Lighting a candle in the dark: advances in genetics and gene therapy of recessive retinal dystrophies. *J Clin Invest*. 2010;120:3042–3053.
3. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *Lancet*. 2006;368:1795–1809.
4. Wang F, Wang H, Tuan HF, et al. Next generation sequencing-based molecular diagnosis of retinitis pigmentosa: identification of a novel genotype-phenotype correlation and clinical refinements. *Hum Genet*. 2014;133:331–345.
5. Daiger SP, Sullivan LS, Bowne SJ. Genes and mutations causing retinitis pigmentosa. *Clin Genet*. 2013;84:132–141.
6. Ran X, Cai WJ, Huang XF, et al. 'RetinoGenetics': a comprehensive mutation database for genes related to inherited retinal degeneration. *Database (Oxford)*. 2014; 2014:bau047.
7. Sullivan LS, Bowne SJ, Birch DG, et al. Prevalence of disease-causing mutations in families with autosomal dominant retinitis pigmentosa: a screen of known genes in 200 families. *Invest Ophthalmol Vis Sci*. 2006;47:3052–3064.
8. Xu Y, Guan L, Shen T, et al. Mutations of 60 known causative genes in 157 families with retinitis pigmentosa based on exome sequencing. *Hum Genet*. 2014;133:1255–1271.
9. Perez-Carro R, Corton M, Sanchez-Navarro I, et al. Panel-based NGS reveals novel pathogenic mutations in autosomal recessive retinitis pigmentosa. *Sci Rep*. 2016;6:19531.
10. Patel N, Aldahmesh MA, Alkuraya H, et al. Expanding the clinical, allelic, and locus heterogeneity of retinal dystrophies. *Genet Med*. 2016;18:554–562.
11. Beryozkin A, Shevah E, Kimchi A, et al. Whole exome sequencing reveals mutations in known retinal disease genes in 33 out of 68 Israeli families with inherited retinopathies. *Sci Rep*. 2015;5:13187.
12. Maranhao B, Biswas P, Gottsch ADH, et al. Investigating the molecular basis of retinal degeneration in a familial cohort of Pakistani decent by exome sequencing. *PLoS One*. 2015;10: e0136561.
13. Weisschuh N, Mayer AK, Strom TM, et al. Mutation detection in patients with retinal dystrophies using targeted next generation sequencing. *PLoS One*. 2016;11:e0145951.
14. Abu-Safieh L, Alrashed M, Anazi S, et al. Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. *Genome Res*. 2013;23:236–247.
15. Huang XF, Huang F, Wu KC, et al. Genotype-phenotype correlation and mutation spectrum in a large cohort of patients with inherited retinal dystrophy revealed by next-generation sequencing. *Genet Med*. 2015;17:271–278.
16. Ellingford JM, Barton S, Bhaskar S, et al. Molecular findings from 537 individuals with inherited retinal disease [published online ahead of print May 11, 2016]. *J Med Genet*. doi:10.1136/jmedgenet-2016-103837.
17. Eisenberger T, Neuhaus C, Khan AO, et al. Increasing the yield in targeted next-generation sequencing by implicating CNV analysis, non-coding exons and the overall variant load: the example of retinal dystrophies. *PLoS One*. 2013;8: e78496.
18. Lindstrand A, Davis EE, Carvalho CM, et al. Recurrent CNVs and SNVs at the NPHP1 locus contribute pathogenic alleles to Bardet-Biedl syndrome. *Am J Hum Genet*. 2014;94:745–754.
19. Coppieters F, Todeschini AL, Fujimaki T, et al. Hidden genetic variation in LCA9-associated congenital blindness explained by 5′UTR mutations and copy-number variations of NMNAT1. *Hum Mutat*. 2015;36:1188–1196.
20. AlMoallem B, Bauwens M, Walraedt S, et al. Novel FRMD7 mutations and genomic rearrangement expand the molecular pathogenesis of X-linked idiopathic infantile nystagmus. *Invest Ophthalmol Vis Sci*. 2015;56:1701–1710.
21. Blanco-Kelly F, Jaijo T, Aller E, et al. Clinical aspects of Usher syndrome and the USH2A gene in a cohort of 433 patients. *JAMA Ophthalmol*. 2015;133:157–164.
22. Coppieters F, Casteels I, Meire F, et al. Genetic screening of LCA in Belgium: predominance of CEP290 and identification of potential modifier alleles in AHI1 of CEP290-related phenotypes. *Hum Mutat*. 2010;31:E1709–E1766.
23. Kaeser PS, Deng L, Fan M, Sudhof TC. RIM genes differentially contribute to organizing presynaptic release sites. *Proc Natl Acad Sci U S A*. 2012;109:11830–11835.
24. Johnson S, Halford S, Morris AG, et al. Genomic organisation and alternative splicing of human RIM1, a gene implicated in autosomal dominant cone-rod dystrophy (CORD7). *Genomics*. 2003;81:304–314.
25. Khateb S, Hanany M, Khalaileh A, et al. Identification of genomic deletions causing inherited retinal degenerations by coverage analysis of whole exome sequencing data. *J Med Genet*. 2016;53:600–607.

Investigative Ophthalmology & Visual Science