

# Comprehensive Molecular Diagnosis of a Large Cohort of Japanese Retinitis Pigmentosa and Usher Syndrome Patients by Next-Generation Sequencing

Maho Oishi,<sup>1</sup> Akio Oishi,<sup>1</sup> Norimoto Gotoh,<sup>2</sup> Ken Ogino,<sup>1</sup> Koichiro Higasa,<sup>2</sup> Kei Iida,<sup>3</sup> Yukiko Makiyama,<sup>1</sup> Satoshi Morooka,<sup>1</sup> Fumihiko Matsuda,<sup>2</sup> and Nagahisa Yoshimura<sup>1</sup>

<sup>1</sup>Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan

<sup>2</sup>Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

<sup>3</sup>Medical Research Support Center, Kyoto University Graduate School of Medicine, Kyoto, Japan

Correspondence: Akio Oishi, 54 Kawahara, Shogoin, Sakyo, Kyoto 606-8507, Japan; aquio@kuhp.kyoto-u.ac.jp.

Submitted: August 13, 2014

Accepted: October 7, 2014

Citation: Oishi M, Oishi A, Gotoh N, et al. Comprehensive molecular diagnosis of a large cohort of Japanese retinitis pigmentosa and Usher syndrome patients by next-generation sequencing. *Invest Ophthalmol Vis Sci.* 2014;55:7369–7375. DOI: 10.1167/iovs.14-15458

**PURPOSE.** Retinitis pigmentosa (RP), a major cause of blindness in developed countries, has multiple causative genes; its prevalence differs by ethnicity. Usher syndrome is the most common form of syndromic RP and is accompanied by hearing impairment. Although molecular diagnosis is challenging, recent technological advances such as targeted high-throughput resequencing are efficient screening tools.

**METHODS.** We performed comprehensive molecular testing in 329 Japanese RP and Usher syndrome patients by using a custom capture panel that covered the coding exons and exon/intron boundaries of all 193 known inherited eye disease genes combined with Illumina HiSeq 2500. Candidate variants were screened using systematic data analyses, and their potential pathogenicity was assessed according to the frequency of the variants in normal populations, in silico prediction tools, and compatibility with known phenotypes or inheritance patterns.

**RESULTS.** Molecular diagnoses were made in 115/317 RP patients (36.3%) and 6/12 Usher syndrome patients (50%). We identified 104 distinct mutations, including 66 novel mutations. *EYS*, *USH2A*, and *RHO* were common causative genes. In particular, mutations in *EYS* accounted for 15.0% of the autosomal recessive/simplex RP patients or 10.7% of the entire RP cohort. Among the 189 previously reported mutations detected in the current study, 55 (29.1%) were found commonly in Japanese or other public databases and were excluded from molecular diagnoses.

**CONCLUSIONS.** By screening a large cohort of patients, this study catalogued the genetic variations involved in RP and Usher syndrome in a Japanese population and highlighted the different distribution of causative genes among populations.

**Keywords:** retinitis pigmentosa, Usher syndrome, next-generation sequencing, targeted resequencing

Retinitis pigmentosa (RP) is the most frequent subtype of inherited retinal degeneration, and its prevalence is 1 in 3000 to 5000 individuals worldwide.<sup>1</sup> Retinitis pigmentosa is clinically and genetically heterogeneous. For example, RP accompanied by hearing impairment is termed Usher syndrome, which is the most common form of syndromic RP. To date, mutations in 73 genes are known to be responsible for RP or Usher syndrome. In addition, 141 genes are associated with other subtypes of inherited retinal diseases (RetNet, <https://sph.uth.edu/retnet/> [in the public domain]). However, no known single gene mutation accounts for more than 10% of RP patients in any reported studies.<sup>2</sup> Although the screening of known genes should be the first step for identifying novel causative genes, sequencing all these genes in individual patients has been virtually impossible.

In principle, the causative gene cannot be predicted from the phenotype. Inheritance patterns, which include autosomal dominant (ad), autosomal recessive (ar), X-linked (xl), digenic,<sup>3</sup> and maternal or mitochondrial,<sup>4,5</sup> can help to narrow down the

number of candidate genes. However, the inheritance pattern cannot always be determined because of patients' insufficient memory or lack of information regarding their family history. In addition, mutations in the same gene sometimes cause different phenotypes or different inheritance traits, which make examinations complicated.

The most widely used methods for the molecular diagnosis of RP, such as Sanger sequencing and arrayed primer extension (APEX) chips (Asper Ophthalmics, Tartu, Estonia), have limitations. The Sanger technique can determine sequences accurately; however, it is labor-intensive, time-consuming, and cost-prohibitive for screening multiple genes. Arrayed primer extension is designed to detect only selected known mutations; therefore, a novel mutation cannot be detected. Specifically, the detection rate of APEX is <10% to 15%,<sup>6</sup> or even lower, depending on the ethnicity of the patients.<sup>7</sup>

Recently, the development of next-generation sequencing (NGS) technology enabled the sequencing of all exons or even the whole genome of an individual. The technology had been

applied to RP/Usher syndrome, and several studies reported a significantly higher diagnosis rate (25%–55%) compared with that achieved using conventional methods.<sup>8–14</sup> Furthermore, the results revealed a different prevalence of causative genes among different ethnicities. For example, studies conducted in North America or China did not report patients with mutations in *EYS*,<sup>11,12</sup> which comprise ~20% of arRP cases in Japan,<sup>15,16</sup> or with mutations in *DHDDS*, which account for ~10% of arRP cases in Ashkenazi Jewish populations.<sup>17,18</sup> These results highlight the importance of making specific genetic catalogues for each ethnicity.

In the current study, we performed a comprehensive molecular analysis of 329 Japanese RP or Usher syndrome patients. We analyzed all retinal and optic nerve disease genes reported in the RetNet database at the time of designing this study (193 genes). In addition, in an attempt to identify novel causative genes, an additional 172 genes were also analyzed using a custom-designed targeted resequencing technique.

## METHODS

All procedures used in this study adhered to the tenets of the Declaration of Helsinki. The institutional review boards and the ethics committees of each institution approved the study protocols. All patients and their relatives were fully informed of the purpose and procedures of this study, and written consent was obtained from each participant.

### Study Subjects

We examined 329 Japanese patients with RP who visited the Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan, between January 2011 and December 2012 and agreed to participate in the study. Among these, 12 patients exhibited hearing impairment and were diagnosed with Usher syndrome. An additional 26 cases with various forms of inherited retinal diseases, carrying known variants, were investigated to evaluate the integrity of the current approach.

Retina specialists made all diagnoses of RP based on comprehensive ophthalmologic examinations, and disorder of hearing and/or equilibrium was confirmed by otolaryngologists in Usher syndrome cases. All patients underwent visual acuity measurements, slit-lamp biomicroscopy, ophthalmoscopy, fundus photography, optical coherence tomography, Goldman visual field testing, and electroretinography, according to the protocol of the International Society for Clinical Electrophysiology of Vision.<sup>19</sup> Wide-field fundus autofluorescence<sup>20</sup> was also available for most patients. Pedigrees were constructed based on patient interviews. A peripheral blood sample was taken from every patient and his or her family members when available. Genomic DNA was extracted from the peripheral blood by using a DNA extraction kit (Quick-Gene-610L; Fujifilm, Minato, Tokyo, Japan). The quantity and quality of the DNA were verified using a dsDNA HS Assay kit on a Qubit (Life Technologies, Carlsbad, CA, USA) and a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA).

### Target Capture and Next-Generation Sequencing

Target regions were captured using reagents from a custom-design HaloPlex Target Enrichment kit 2.5 Mb (Agilent Technologies, Santa Clara, CA, USA), according to the HaloPlex Target Enrichment System Automation Protocol for Illumina Sequencing Version D.2. SureDesign software (Agilent Technologies) was used to design the custom HaloPlex

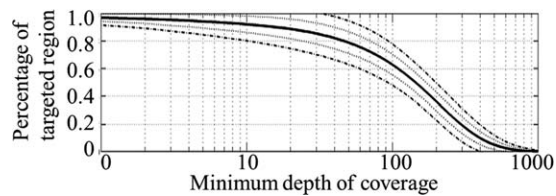
capture assay. The capture panel consisted of 2,433,298 base pairs (bp) covering all coding and noncoding exons with flanking exon/intron boundaries ( $\pm 25$  bp) of 365 genes, including all 193 retinal and optical disease genes that had been reported in RetNet at the time of designing this study (December 26, 2012), as well as the 172 genes expressed at the highest level in rod or cone photoreceptors (Supplementary Table S1).<sup>21</sup> Precapture Illumina libraries were then generated according to the manufacturer's protocol. Briefly, the protocol consisted of the following four steps: (1) digestion of 225 ng genomic DNA in eight different restriction reactions; (2) hybridization of the digested fragments to probes whose ends are complementary to the target fragments (during hybridization, the fragments were circularized and sequencing motifs, including index sequences, were incorporated); (3) capture of the target DNA using streptavidin beads followed by ligation of the circularized fragments; (4) PCR amplification of the captured target libraries, followed by purification and adaptor-dimer removal by using AMPure XP beads (Beckman Coulter, Brea, CA, USA). For each capture reaction, 47 or 48 libraries were pooled. The concentration of the pooled libraries was determined using a 2100 Bioanalyzer (Agilent Technologies). The pooled libraries were quantified and sequenced on the Illumina HiSeq2500 as 100-bp paired-end reads following the manufacturer's protocols.

### Data Analysis

Sequence reads were aligned to the reference genome (National Center for Biotechnology Information [NCBI] Build 37) by using the Burrows-Wheeler Aligner<sup>22</sup> after trimming the adapter sequence (AGATCG) using custom Perl script. After recalibration of base quality values and local realignment, single nucleotide variant and insertion/deletion (indel) calling was performed using the Genome Analysis Toolkit.<sup>23</sup> To filter out common SNPs and indels, dbSNP (Build 138) (NCBI, <http://www.ncbi.nlm.nih.gov/SNP/> [in the public domain]), 1000 Genomes,<sup>24</sup> National Heart, Lung and Blood Institute GO Exome Sequencing Project (ESP6500, <http://evs.gs.washington.edu/EVS/> [in the public domain]), and the Human Genetic Variation Database (HGVD; <http://www.genome.med.kyoto-u.ac.jp/SnpDB/> [in the public domain]) databases were used. The HGVD contains genetic variations determined using the exome sequencing of 1208 Japanese control subjects. Variant annotation was performed using ANNOVAR.<sup>25</sup> To assess the pathogenicity of novel missense variants, five types of prediction scores for amino acid substitutions (SIFT, Polyphen2, LRT, MutationTaster, and MutationAssessor) and two conservation scores (PhyloP and GERP++) obtained from dbNSFP<sup>26</sup> were used.

### Interpretation of the Genetic Variants

Several criteria were applied to select variants for further analysis and to filter out putative false positives caused by alignment artifacts. Analysis focused on variants of the 193 genes reported in RetNet database, which includes all known retinal disease genes. First, variants were selected within coding exons or intronic variants  $\pm 25$  bp from the exon boundaries. Second, variants that covered  $< 10\times$  were excluded. Third, synonymous variants and nonframeshift indels were excluded unless they were listed in the Human Gene Mutation Database (HGMD; BIOBASE, Wolfenbüttel, Lower Saxony, Germany) or were reported previously as being pathogenic alterations. However, it should be noted that this step might filter out true pathogenic variants.



**FIGURE 1.** The percentage of the targeted regions (*y*-axis) that achieved a certain minimum coverage (*x*-axis). The *solid curve* represents the mean of 329 samples, whereas the *dotted* and *dashed lines* indicate one and two standard deviations, respectively. The data reveal that 92.2% of the targeted regions are covered by  $>10\times$ .

### Determination of the Pathogenicity of the Variants

A variant was classified as pathogenic if the following criteria applied. (1) Because RP is a rare Mendelian disease, variants were excluded that had an allele frequency  $> 0.5\%$  (for recessive variants) or  $>0.1\%$  (for dominant variants) in any of the 1000 Genomes database, ESP6500 database, or HGVD. (2) Mutations listed in the HGMD or those identified as pathogenic alterations in previous publications were regarded as pathogenic. (3) Nonsense and frameshift variants were also considered as pathogenic. (4) For a novel missense variant, in silico prediction programs were used to predict its pathogenicity. Only novel missense variants that were predicted to be pathogenic by at least five of seven well-established algorithms were reported. Variants that were predicted to be pathogenic by at least three of five missense prediction programs (SIFT, Polyphen2, LRT, MutationTaster, and Mutation-Assessor) and whose evolutionary conservation scores were  $>0$  in both PhyloP and GERP++ were considered to be pathogenic. (5) For splice-site variants, the prediction program MaxEntScan was used, and these were considered pathogenic if the score differed by  $>5$  between the wild-type and mutated sequences.<sup>27</sup> (6) Variants were adopted that matched the patients' phenotype and the reported inheritance pattern of the respective genes.

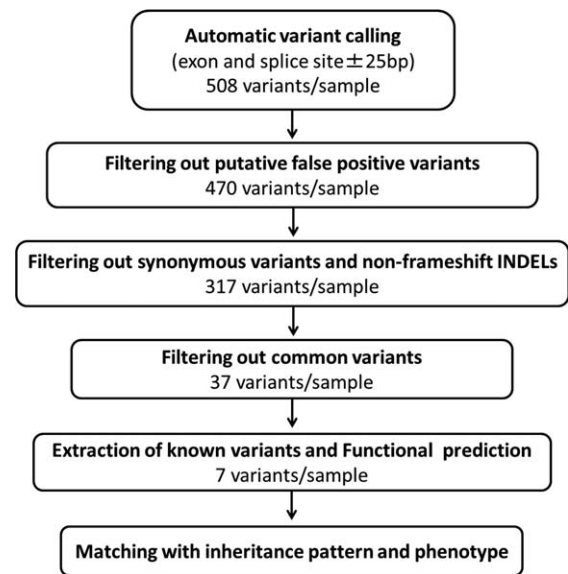
Pathogenic variants in the dominant genes found in simplex cases were regarded as disease-causing mutations only when they were published previously or were confirmed to be a de novo mutation by using parental testing. Variants with a frequency more than that mentioned in criterion 1 were excluded even if they were listed in the HGMD or in previously published reports. All mutations and potential pathogenic variants detected using NGS were validated using conventional Sanger sequencing. Sequencing was performed using an Applied Biosystems (ABI) 3130xl Genetic Analyzer (Life Technologies). Segregation analysis was performed if DNA from family members was available.

### Evaluation of the Sensitivity of the Method

Twenty-six cases carrying 33 known variants in various genes were included as positive controls to evaluate the integrity of the current approach (Supplementary Table S2). All variants were detected using the APEX chip and confirmed using Sanger sequencing before the current study was conducted.

## RESULTS

A total of 329 patients, including 77 adRP, 73 arRP, 6 xlRP, 161 simplex RP, and 12 Usher syndrome patients, were studied. Among these, two pairs of adRP cases and one pair of arRP cases had a blood relationship (two pairs of adRP cases had a parent-child relationship, whereas one pair of arRP cases comprised siblings).



**FIGURE 2.** The data analysis pipeline used to filter and identify the putative pathogenic mutations in this study. Three public databases and seven in silico programs were used to evaluate the pathogenicity of variants. This stringent filtering should have excluded false-positive variants.

### Targeted Sequencing and Data Processing

All exons and the regions 25 bp upstream and downstream of 365 genes were enriched using targeted capture and sequenced using NGS. For each sample, a mean of 7,548,920 reads was generated, approximately 82.2% of which were mapped to the targeted regions. Although a small number of samples (74/329, 22.5%) had relatively low coverage compared with the others (Supplementary Fig. S1), the mean and median coverage of the 193 RetNet genes in all samples was  $250\times$  and  $242\times$ , respectively. Within the targeted region, 92.2% of the bases had coverage  $> 10\times$ , and 88.7% of the bases had coverage  $> 20\times$ , suggesting that sufficient coverage was achieved (Fig. 1). Among the 3619 targeted exons, only 1.58% covered  $< 5\times$  (Supplementary Table S3).

For each sample, a mean of 1577 raw variants and small indels were identified initially by using automated variant detection. After filtering out putative false-positive variants and synonymous or nonframeshift variants, a mean of 317 variants remained. After excluding common variants in any of the variant databases (as described in Materials and Methods section on interpretation of genetic variants), functional prediction tools were applied, and a mean 7.08 rare variants that were likely to cause a deleterious protein coding change were detected per sample (Fig. 2).

### Identification of Mutations in Positive Controls

All 33 variants present in the 26 positive control samples were detected, in spite of stringent filtering steps taken to exclude putative false-positive variants. This indicates that the combination of HaloPlex and Illumina HiSeq resulted in high-sensitivity sequencing and confirmed validity of the data processing methods (Supplementary Table S2).

### Identification of Pathogenic Mutations in 329 RP or Usher Syndrome Patients

Mutations listed in the HGMD or in previous publications as causing retinal diseases, as well as novel pathogenic variants



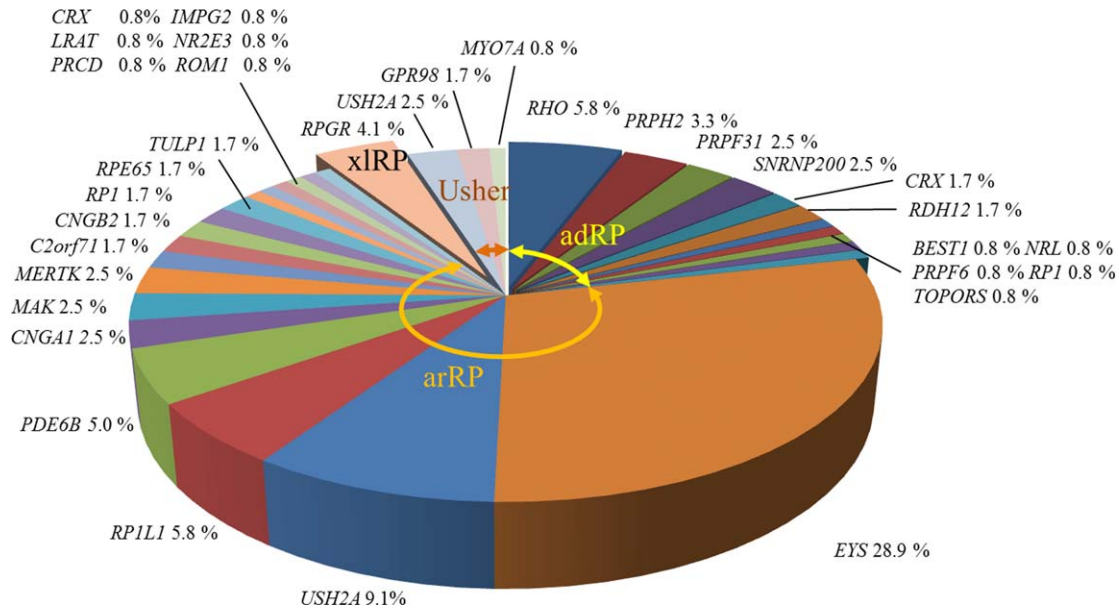


FIGURE 3. The percentage of causative genes with each inheritance trait that were identified in 115 RP patients and six Usher syndrome patients. Mutations in *EYS* and *USH2A* were common in the population.

that were predicted to be pathogenic after the stringent filtering steps, were identified. Specifically, pathogenic mutations were identified for 115/317 (36.3%) RP patients and 6/12 (50%) Usher syndrome patients. A total of 103 distinct mutations were identified, including a large number of novel mutations ( $n = 65$ ). The novel mutations consisted of 23 missense (35.4%), 19 frameshift (29.2%), 17 nonsense (26.2%), and six splice-site (9.2%) mutations (Supplementary Tables S4–S9). The pathogenicity of the novel missense mutations was supported by a combination of in silico prediction programs (Supplementary Table S8). A summary of the causative genes is shown in Figure 3 and the Table.

**Patients Carrying Mutations in Known RP Genes.** We identified 121 patients who carried pathogenic mutations in known RP or Usher syndrome genes (Supplementary Tables S4–S6). The diagnostic efficiency of each hereditary form was 27/77 (35.1%) in adRP, 37/73 (50.7%) in arRP, 1/6 (16.7%) in xLRP, 50/161 (31.1%) in simplex RP, and 6/12 (50%) in Usher syndrome. Among the 50 simplex cases whose mutations were identified, 3 were due to mutations in ad genes (K6413, K6210, and K6419). One of the 77 ad cases and 3 of 161 simplex cases turned out to be xl with mutations in *RPGR* (K1881, K6101, K6170, and K6292). Three cases initially diagnosed as ad were determined to be ar with known mutations in *EYS* (K6006, K6105, and K6195).

We identified 48 patients carrying mutations that are known to cause RP or Usher syndrome (Supplementary Table S4). Among these cases diagnosed with high confidence, 24/48 (50.0%) patients carried mutations in the *EYS* gene. Proband K1908 and proband K6184 had three mutations, including a novel nonsense mutation c.1750G>T in *EYS* or a novel missense mutation c.2653C>T in *USH2A*, respectively.

Forty-five patients carried at least one novel loss-of-function (LOF) mutation (nonsense, frameshift, or splicing mutations) in known RP or Usher syndrome genes (Supplementary Table S5). Among this group diagnosed with relatively high confidence, 44 distinct pathogenic mutations were identified, including 6 previously reported mutations. One pair of familial cases (K6001 and K6003) was included in this group.

In addition, 28 patients who carried one or more novel putative missense mutations in RP or Usher syndrome genes

TABLE. List of the Causative Genes in 121 RP and Usher Syndrome Patients Identified in the Current Study

Gene	<i>n</i>
RP	
<i>EYS</i>	35
<i>USH2A</i>	11
<i>RHO</i>	7
<i>RPILI</i>	7
<i>PDE6B</i>	6
<i>RPGR</i>	5
<i>PRPH2</i>	4
<i>CNGA1</i>	3
<i>CRX</i>	3
<i>MAK</i>	3
<i>MERTK</i>	3
<i>PRPF31</i>	3
<i>RP1</i>	3
<i>SNRNP200</i>	3
<i>C2orf71</i>	2
<i>CNGB1</i>	2
<i>RDH12</i>	2
<i>RPE65</i>	2
<i>TULP1</i>	2
<i>BEST1</i>	1
<i>IMPG2</i>	1
<i>LRAT</i>	1
<i>NR2E3</i>	1
<i>NRL</i>	1
<i>PRCD</i>	1
<i>PRPF6</i>	1
<i>ROM1</i>	1
<i>TOPORS</i>	1
Usher syndrome	
<i>USH2A</i>	3
<i>GPR98</i>	2
<i>MYO7A</i>	1
Total	121

were identified (Supplementary Table S6). In this group with lower confidence diagnosis, 32 distinct pathogenic mutations were identified, including 3 previously reported mutations, 3 novel LOF mutations, and 26 novel missense mutations in 17 genes.

**Patients Carrying Known Mutations in Other Retinal Disease Genes.** Two patients carried reported mutations that are known to cause other retinal diseases (Supplementary Table S7). Proband K6043 carried c.1511A>G in *JAG1*, which causes Alagille syndrome. Proband K1956 carried c.808C>T in *CHM*, which causes choroideraemia.

### Revision of the Initial Clinical Diagnosis

The clinical symptoms of all patients identified as carrying known mutations in other retinal diseases were carefully reassessed. In the group with lower confidence diagnoses, three patients (K6266, K6274, and K6348) carried mutations known to cause Best macular dystrophy, pattern dystrophy, or retinitis punctata albescens, respectively (Supplementary Table S6). Although these mutations have been reported to cause other retinal diseases, their clinical characteristics were typical of RP; therefore, these variants were classified as novel missense mutations that cause RP. All of these were variants in known RP genes, and their pathogenicity was supported by *in silico* prediction tools (Supplementary Table S8). In an additional two patients carrying known mutations in other retinal disease genes (Supplementary Table S7), the clinical diagnosis of xl case K1956, who carried a reported hemizygous nonsense mutation known to cause choroideraemia, was revised to choroideraemia. Choroideraemia is a nonsyndromic choroidal and retinal disease that presents with symptoms and clinical features similar to those of advanced RP. In contrast, the simplex case K6043, who carries a mutation known to cause Alagille syndrome, remains unsolved because of the lack of characteristic systemic disorders.

### Reported Mutations With a High Allele Frequency in Public Databases

Using filtering and annotation procedures, some of the reported mutations were revealed to be common in healthy individuals. Because RP is a rare Mendelian disease, we considered that variants that had an allele frequency > 0.5% (for recessive variants) or >0.1% (for dominant variants) in any of the public databases were pathogenic in the current study. After these analyses, 55 of the 189 reported mutations known to cause retinal or optic nerve diseases were excluded (Supplementary Table S10).

## DISCUSSION

In the present study, we screened 329 Japanese RP or Usher syndrome patients, and made molecular diagnoses in 122 cases (37.1%). To our knowledge, the present cohort is the largest among those in studies using NGS for the comprehensive molecular diagnosis of RP. In addition, this was the first such study in Japanese individuals; the data confirmed the high prevalence of mutations in *EYS* and *USH2A* gene among this population.

The present study revealed that the prevalence of causative genes differed in Japanese populations compared with that in other ethnicities. The frequency of *EYS* mutations was up to 11% (35/317) in the Japanese RP population, which is the highest percentage by which a single gene accounts for RP.<sup>2</sup> The association between *EYS* mutations and RP was first reported in 2008.<sup>28</sup> Subsequently, several reports revealed a

prevalence of *EYS* mutations of 5% to 10% of European arRP patients,<sup>13,29,30</sup> whereas the prevalence of *EYS*-associated RP was as high as 20% in Japanese arRP patients.<sup>15,16</sup> The present study confirmed the significant influence of *EYS* in Japanese populations. Specifically, the allele frequency of c.8805C>A and c.4957dupA, the most frequent mutations in *EYS*, in the control Japanese population was 0.34% and 0.32%, respectively. This high frequency of carriers accounts for the detection of *EYS* mutations in three probands who were presumed to be arRP.

The present results have some implications regarding a role for the *RP1L1* gene in RP. Homozygous mutations in *RP1L1* were found in seven patients, which account for 7.8% of the arRP cases. Although mutations in this gene were first reported in patients with ad occult macular dystrophy,<sup>31</sup> homozygous mutations of *RP1L1* were reported subsequently in patients with arRP.<sup>32</sup> The prevalence of *RP1L1* mutations in the current study was considerably higher than that reported in the previous study; Davidson et al.<sup>32</sup> revealed one patient by using homozygosity mapping, and only one other patient among 285 RP patients. Differences in the mutation allele frequency among ethnicities might also be observed in *RP1L1*. Because *RP1L1* interacts with *RP1*, modifications to *RP1*-associated RP were also suggested.<sup>33</sup> However, none of the patients with putative disease-causing mutations in *RP1L1* carried mutations in *RP1* gene in the current study or in a previous report.<sup>32</sup> Therefore, the hypothesis that *RP1* and *RP1L1* cause digenic RP is yet to be confirmed.

Before the advent of NGS, screening using denaturing high-performance liquid chromatography (dHPLC) or high-resolution melting analysis followed by Sanger sequence was used to detect causative mutations in genetically heterogeneous diseases such as RP. Jin et al.<sup>34</sup> applied dHPLC and Sanger sequencing to Japanese RP patients and detected causative mutations in 14.1% of cases. The difference in the prevalence of causative mutations among ethnicities could partially explain the unsatisfactory detection rate in this previous report. The current study revealed the highest detection rate in Japanese populations to date, which was comparable to that in populations of other ethnicities.<sup>8-14</sup> These results confirm that comprehensive screening using NGS is an effective strategy, independent of the patient cohort.

The present results suggest that some of the previously reported mutations were not disease causing. In the present study, we used a public genome database containing 1208 healthy Japanese individuals, as well as the 1000 Genomes and ESP6500 databases to screen the candidate variants. This revealed that some mutations registered in the HGMD are common in Japanese populations, suggesting that these mutations were benign variants. Therefore, information regarding disease-causing mutations should be revised based on the current evidence.

We employed targeted exome resequencing rather than whole-exome sequencing (WES). Although WES could become a standard screening method in the near future, targeted exome resequencing retains some advantages. Specifically, it can achieve a higher coverage rate in the region of interest, and can screen more patients in a single assay, thereby lowering cost.<sup>35</sup> As such, targeted exome resequencing is a reasonable option to screen highly heterogeneous groups of diseases such as RP.

Although molecular diagnoses were made in 37.7% of the cases in the current study, causative mutations are yet to be elucidated in the remaining 62.3% of the patients. The technical and methodological limitations of our approach might be one of the reasons for this. We confirmed the sensitivity of the method using 26 patients carrying 33 variants identified previously.<sup>7</sup> Next-generation sequencing detected all

these variants successfully, which verified the high sensitivity of the current approach. However, NGS technology is not good at reading certain specific sequences including GC-rich regions, repeated sequences, copy-number variations, and large deletions. Intronic mutations, synonymous mutations, and nonframeshift mutations were also not covered by the current method. In addition, some cases resulted in relatively low coverage compared with the other samples. This might have been caused by technical challenges during the generation of precapture libraries. Although the mean coverage achieved was 244×, which seemed to be sufficient for the identification of variants, the percentage of bases covered >10× was 81%, and variants with low coverage were excluded (Supplementary Fig. S1). A more important reason for the lack of detection is insufficient knowledge of RP-causative genes. Although more than 60 causative genes have been identified to date, this number is still increasing. Therefore, further exploration of novel genes is required to achieve an optimal detection rate.

In conclusion, the present study screened the largest sample of Japanese RP patients to date and described the genetic catalogue for the cohort. The data confirmed differences in the RP-causative genes among ethnicities, and highlighted the importance of integrating studies from multiple populations for a deeper understanding of the disease-causing mutations. We are currently identifying novel genes in the captured candidates. Future studies including WES or even whole-genome sequencing in various ethnicities might boost the identification of novel causative genes of RP.

### Acknowledgments

We thank Jun Zhu, PhD; Yoshiyuki Wakabayashi, PhD; and Yanqin Yang, MD, MSc, at DNA Sequencing and Computational Biology Core, National Heart, Lung and Blood Institute, National Institutes of Health, for their advice on technical issues and invaluable comments to improve the manuscript. DNA sequencing analysis with the Sanger method was performed at the Medical Research Support Center, Graduate School of Medicine, Kyoto University (chief director, Masatoshi Hagiwara, MD, PhD). We thank Hatsue Hamanaka for her technical support in the preparation and sequencing of the genetic samples.

Supported in part by the Japan Ministry of Health, Labour and Welfare (No. 12103069), a grant-in-aid for scientific research (No. 26861445), and the Japanese Retinitis Pigmentosa Society. The authors alone are responsible for the content and writing of the paper.

Disclosure: **M. Oishi**, None; **A. Oishi**, None; **N. Gotoh**, None; **K. Ogino**, None; **K. Higasa**, None; **K. Iida**, None; **Y. Makiyama**, None; **S. Morooka**, None; **F. Matsuda**, None; **N. Yoshimura**, None

### References

- Chizzolini M, Galan A, Milan E, Sebastiani A, Costagliola C, Parmeggiani F. Good epidemiologic practice in retinitis pigmentosa: from phenotyping to biobanking. *Curr Genomics*. 2011;12:260–266.
- Wang DY, Chan WM, Tam PO, et al. Gene mutations in retinitis pigmentosa and their clinical implications. *Clin Chim Acta*. 2005;351:5–16.
- Kajiwaru K, Berson EL, Dryja TP. Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. *Science*. 1994;264:1604–1608.
- Lestienne P, Ponsot G. Kearns-Sayre syndrome with muscle mitochondrial DNA deletion. *Lancet*. 1988;1:885.

- Moraes CT, DiMauro S, Zeviani M, et al. Mitochondrial DNA deletions in progressive external ophthalmoplegia and Kearns-Sayre syndrome. *N Engl J Med*. 1989;320:1293–1299.
- Avila-Fernandez A, Cantalapiedra D, Aller E, et al. Mutation analysis of 272 Spanish families affected by autosomal recessive retinitis pigmentosa using a genotyping microarray. *Mol Vis*. 2010;16:2550–2558.
- Ogino K, Oishi A, Makiyama Y, et al. Genotype screening of retinal dystrophies in the Japanese population using a microarray [in Japanese]. *Nippon Ganka Gakkai Zasshi*. 2013;117:12–18.
- Simpson DA, Clark GR, Alexander S, Silvestri G, Willoughby CE. Molecular diagnosis for heterogeneous genetic diseases with targeted high-throughput DNA sequencing applied to retinitis pigmentosa. *J Med Genet*. 2011;48:145–151.
- Neveling K, Collin RW, Gilissen C, et al. Next-generation genetic testing for retinitis pigmentosa. *Hum Mutat*. 2012;33:963–972.
- Shanks ME, Downes SM, Copley RR, et al. Next-generation sequencing (NGS) as a diagnostic tool for retinal degeneration reveals a much higher detection rate in early-onset disease. *Eur J Hum Genet*. 2013;21:274–280.
- Fu Q, Wang F, Wang H, et al. Next-generation sequencing-based molecular diagnosis of a Chinese patient cohort with autosomal recessive retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2013;54:4158–4166.
- 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.
- Glockle N, Kohl S, Mohr J, et al. Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. *Eur J Hum Genet*. 2014;22:99–104.
- 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.
- Hosono K, Ishigami C, Takahashi M, et al. Two novel mutations in the EYS gene are possible major causes of autosomal recessive retinitis pigmentosa in the Japanese population. *PLoS One*. 2012;7:e31036.
- Iwanami M, Oshikawa M, Nishida T, Nakadomari S, Kato S. High prevalence of mutations in the EYS gene in Japanese patients with autosomal recessive retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2012;53:1033–1040.
- Zelinger L, Banin E, Obolensky A, et al. A missense mutation in DHDDS, encoding dehydrodolichyl diphosphate synthase, is associated with autosomal-recessive retinitis pigmentosa in Ashkenazi Jews. *Am J Hum Genet*. 2011;88:207–215.
- Zuchner S, Dallman J, Wen R, et al. Whole-exome sequencing links a variant in DHDDS to retinitis pigmentosa. *Am J Hum Genet*. 2011;88:201–206.
- Marmor MF, Fulton AB, Holder GE, Miyake Y, Brigell M, Bach M. ISCEV standard for full-field clinical electroretinography (2008 update). *Doc Ophthalmol*. 2009;118:69–77.
- Oishi A, Ogino K, Makiyama Y, Nakagawa S, Kurimoto M, Yoshimura N. Wide-field fundus autofluorescence imaging of retinitis pigmentosa. *Ophthalmology*. 2013;120:1827–1834.
- Brooks MJ, Rajasimha HK, Roger JE, Swaroop A. Next-generation sequencing facilitates quantitative analysis of wild-type and Nrl(-/-) retinal transcriptomes. *Mol Vis*. 2011;17:3034–3054.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–1760.

23. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491-498.
24. Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467:1061-1073.
25. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164.
26. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat.* 2013;34:E2393-E2402.
27. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004;11:377-394.
28. Abd El-Aziz MM, Barragan I, O'Driscoll CA, et al. EYS, encoding an ortholog of *Drosophila* spacemaker, is mutated in autosomal recessive retinitis pigmentosa. *Nat Genet.* 2008;40:1285-1287.
29. Audo I, Sahel J-A, Mohand-Saïd S, et al. EYS is a major gene for rod-cone dystrophies in France. *Hum Mutat.* 2010;31:E1406-E1435.
30. Littink KW, van den Born LI, Koenekoop RK, et al. Mutations in the EYS gene account for approximately 5% of autosomal recessive retinitis pigmentosa and cause a fairly homogeneous phenotype. *Ophthalmology.* 2010;117:2026-2033, 2033 e1-7.
31. Akahori M, Tsunoda K, Miyake Y, et al. Dominant mutations in RP1L1 are responsible for occult macular dystrophy. *Am J Hum Genet.* 2010;87:424-429.
32. Davidson AE, Sergouniotis PI, Mackay DS, et al. RP1L1 variants are associated with a spectrum of inherited retinal diseases including retinitis pigmentosa and occult macular dystrophy. *Hum Mutat.* 2013;34:506-514.
33. Yamashita T, Liu J, Gao J, et al. Essential and synergistic roles of RP1 and RP1L1 in rod photoreceptor axoneme and retinitis pigmentosa. *J Neurosci.* 2009;29:9748-9760.
34. Jin ZB, Mandai M, Yokota T, et al. Identifying pathogenic genetic background of simplex or multiplex retinitis pigmentosa patients: a large scale mutation screening study. *J Med Genet.* 2008;45:465-472.
35. Sikkema-Raddatz B, Johansson LE, de Boer EN, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat.* 2013;34:1035-1042.