Sequencing Arrays for Screening Multiple Genes Associated with Early-Onset Human Retinal Degenerations on a High-Throughput Platform

Md Nawajes A. Mandal,1 John R. Heckenlively,1 Tracy Burch,1 Lianchun Chen,1 Vidyullatha Vasireddy,1 Robert K. Koenekeop,2 Paul A. Sieving,3 and Radha Ayyagari1

PURPOSE. To develop and apply microarray-based resequencing technology to detect sequence alterations in multiple autosomal recessive retinal disease genes on a single high-throughput platform.

METHODS. Oligonucleotides corresponding to both strands of the target exons and the flanking intron sequences of 29,214 bp from 11 genes associated with autosomal recessive retinitis pigmentosa (arRP) were tiled on 20 × 25-μm microarrays (arRP-I arrays). A total of 155 exons were amplified from 35 arRP patient DNA samples, with each sample being sequenced on an arRP-I chip by hybridization.

RESULTS. With the arRP-I arrays, 97.6% of the tiled sequence were determined with more than 99% accuracy and reproducibility. Of the 2.4% unread sequence, 89.5% involved stretches of G or C. In analyzing the 903,140-bp sequence from the 35 patient samples, 506 sequence changes have been detected in which 386 are previously reported alterations, and 120 are novel. In addition to four known causative mutations, six novel sequence changes that are potentially pathogenic were observed. Additional analysis is needed to determine whether these changes are responsible for arRP in these patients.

CONCLUSIONS. The use of microarray for sequencing is a novel approach, and the arRP-I chip is the first successful application of this technology for determining sequence alteration in multiple disease-related genes. These arrays can be used for high-throughput genotyping of patients with relevant retinal conditions. In addition, these arrays offer a unique opportunity to interrogate complex patterns of inheritance due to the involvement of more than one gene by screening multiple genes on a single platform. (Invest Ophthalmol Vis Sci. 2005;46:3355–3362) DOI:10.1167/iovs.05-00007

From the 1Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, Michigan; 2McGill Ocular Genetics Laboratory, McGill University Health Center, MTL Children’s Hospital, Montreal, Quebec, Canada; and the 3National Eye Institute, National Institutes of Health, Bethesda, Maryland.

Supported by National Eye Institute Grant EY13198 (RA); the Foundation Fighting Blindness (RA, JHR); Research to Prevent Blindness Inc. (RA); National Eye Institute Core Grants EY07003 and EY07060 to the University of Michigan, Department of Ophthalmology and Visual Sciences; the Foundation Fighting Blindness, Canada; and the Fonds de la Recherche en Santé Québec (FRSQ).

Submitted for publication January 4, 2005; revised March 16 and April 6, 2005; accepted April 14, 2005.

Disclosure: M.N.A. Mandal, None; J.R. Heckenlively, None; T. Burch, None; L. Chen, None; V. Vasireddy, None; R.K. Koenekeop, None; P.A. Sieving, None; R. Ayyagari, None

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Radha Ayyagari, Department of Ophthalmology and Visual Sciences, University of Michigan, 1000 Wall Street, Room 325, Ann Arbor, MI 48105; ayyagari@umich.edu.

R etinitis pigmentosa (RP) is a heterogeneous group of diseases characterized by progressive retinal degeneration leading to irreversible blindness or severe visual disability. It affects 1 in 3500 individuals worldwide. Clinically, it is characterized by an abnormal electroretinogram (ERG), loss of peripheral vision, and night blindness in the initial stage, which progresses to severe visual disability that can progress to complete vision loss.1 RP can be inherited in an autosomal dominant or recessive or X-linked manner, as well as in digenic, mitochondrial, and simplex modes. Most patients with RP are simplex with no known affected relatives, although some of these may have autosomal or X-linked recessive RP or dominant RP with incomplete penetrance. Despite the nearly indistinguishable clinical characteristics of RP with different modes of inheritance, a broad genetic heterogeneity exists. At least 15 loci for autosomal dominant RP, 18 loci for autosomal recessive RP (arRP), and 5 loci for the X-linked form are known, and additional loci are yet to be mapped.2–4 Phenotypic symptoms of RP can also overlap with various other retinal disease phenotypes.5 In addition, several syndromes involving RP have been described (OMIM; Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/Omim/ provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD). Clinical diagnosis is further complicated by the significant variation in the age of onset, severity, clinical presentation, and penetrance that is reported within and between families with retinal degenerations.6–10 All these issues pose a challenge in providing a definite diagnosis to patients, particularly to those who have no clear family history.

Genes for approximately 110 retinal diseases have been cloned, and significant progress has been made in understanding the mechanism of these degenerations.4 Various strategies to treat these conditions are being actively pursued. Results of studies of treatments for two different forms of early-onset retinal degeneration have been reported to be successful in animal models, and one can anticipate that some of these treatments may become available to patients in the future.11,12 However, most of these treatments are specific to patients with particular genotypes, and thus, determination of the genotype of patients will become essential, for identifying individuals suitable for the available treatments.

The genetic heterogeneity, variable penetrance, and considerable task of analyzing all the known candidate genes temper one’s enthusiasm for genotyping patients with RP. So far, at least 32 genes have been identified for various forms of RP, and many more remain to be cloned. The sequence of all the known arRP genes alone adds up to more than 60 kilobases (kb). Current strategies for screening large numbers of bases for mutations are limited by the availability of high-throughput, sensitive, and rapid mutation-detection technologies.

Sequencing using microarrays, which is more typically referred to as resequencing, offers the promise of determining genotypes with a reasonable amount of effort. The past decade has seen the development of sequencing by hybridization to oligonucleotides on a microarray chip.13–19 This technology
Table 1. arRP Genes Arrayed on the Chips and the Number of Missense Mutations Previously Reported in These Genes in Patients with arRP, LCA or CSNB

<table>
<thead>
<tr>
<th>Genes</th>
<th>Additional Phenotype</th>
<th>Number of Exons</th>
<th>Exons Tiled</th>
<th>Number of arRP Mutations Reported</th>
<th>Number of LCA/CSNB Mutations Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA4</td>
<td></td>
<td>50</td>
<td>50</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>CNGA1</td>
<td></td>
<td>8</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>CRB1</td>
<td>LCA, AD</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>MERTK</td>
<td></td>
<td>19</td>
<td>19</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>PDE6A</td>
<td></td>
<td>22</td>
<td>22</td>
<td>1 (exon 7)</td>
<td>1</td>
</tr>
<tr>
<td>PDE6B</td>
<td>CSNB, AD</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>RGR</td>
<td></td>
<td>7</td>
<td>7</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>RHO</td>
<td></td>
<td>5</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>RLBP1</td>
<td></td>
<td>7</td>
<td>7</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>RPE65</td>
<td>LCA, AR</td>
<td>14</td>
<td>14</td>
<td>15</td>
<td>47</td>
</tr>
<tr>
<td>TULP1</td>
<td></td>
<td>14</td>
<td>13</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>179</td>
<td>155</td>
<td>87</td>
<td>93</td>
</tr>
</tbody>
</table>

has also been commercialized. Recently, significant improvements have been made in this technology, and array platforms that can accommodate a large number of bases have been developed. Using this technology, Maitra et al. have developed MitoChip for detection of human mitochondrial mutations, and Wong et al. used this methodology to screen for mutations in the SARS (severe acute respiratory syndrome) virus. In both these cases, it was a haploid genome that was sequenced. Ours is the first report of the application of these resequencing platforms to analyze a diploid genome for mutations in multiple genes. We have screened 11 arRP genes for sequence changes in the coding regions and the flanking splice sites. Herein, we present the validation and potential application of these arRP-I chips to genotype patients with arRP.

Materials and Methods

Design of Human Recessive RP Genes Resequencing Array

Sequences of ATP-binding cassette, subfamily A (ABCA1), member 4 (ABCA4); cyclic nucleotide gated channel alpha 1 (CNGA1); crumbs homolog 1 (Drosophila; CRB1); C-mer proto-oncogene tyrosine kinase (MERTK); phosphodiesterase 6A, cGMP-specific, rod, alpha (PDE6A); phosphodiesterase 6B, cGMP-specific, rod, beta (PDE6B); retinal G protein coupled receptor (RGR); rhodopsin (RHO); retinaldehyde binding protein 1 (RLBP1); retinal pigment epithelium-specific protein 65KDa (RPE65); and tubby like protein 1 (TULP1) were obtained from the human genome database and used as the reference to select the sequence to be arrayed on the chips. The sequences comprising all coding exons plus 15 bp of flanking intronic sequences were selected to tile on the resequencing array. Repetitive elements and internal duplications that may lead to cross hybridization were identified by using Repeat Masker (http://ftp.genome.washington.edu/cgi-bin/repeatmasker; provided in the public domain by the University of Washington Genome Center, Seattle, WA) and deleted. The selected sequence of the arRP genes was tiled onto the 30-kb DNA sequencing platform available through Affymetrix (Santa Clara, CA). Our custom-designed arRP gene resequencing arrays were fabricated by Affymetrix using standard photolithography and solid-phase DNA synthesis.

Patients and DNA

Seventy individuals from different pedigrees with a clinical diagnosis of arRP were selected for genotyping with the arRP-I chips. Thirty-five of these patients were genotyped with these chips, whereas known mutations from the remaining 35 patients were used for the validation of the performance of these arrays (see Table 4). Affected relatives of 26 of these patients were also available for the study. Informed consent was obtained from each participant according to the procedures of the institutional review boards of participating institutions, and the research adhered to the tenets of the Declaration of Helsinki. Blood samples were obtained, and DNA was isolated from lymphocytes by using standard protocols.

PCR Amplification

Amplification of the 155 selected exons of the 11 arRP genes was performed using primers and conditions described earlier (Table 1). A total of 159 PCR reactions were required for each sample, as more than one reaction is needed to amplify some of the larger exons. In some cases, closely spaced exons were amplified in a single reaction. Depending on the conditions required for particular exons, touchdown or normal PCR cycling was used. In touchdown PCR, the annealing temperature of each reaction cycle is decreased by 0.5°C for 26 cycles, so that the annealing temperature is decreased from 68°C to a touchdown at 55°C in 26 cycles. An additional 14 cycles are performed at 55°C annealing temperature to complete the amplification. As a control for PCR amplification and chip hybridization, a 1.0-kb plasmid DNA (TAG IQ-EX) was amplified in a parallel reaction with the samples, using primers and template included in the control kit (CustomSeq: Affymetrix). The quality and the size of amplicons were verified by agarose gel electrophoresis.

Quantitation and Pooling of the PCR Products

The concentration of amplicons was determined by adding a fluorescent label (Picogreen; Molecular Probes, Eugene, OR) to the samples and measuring fluorescence (Spectro Fluor Plus; Tecan Austria Gesellschaft MBH, Salzburg, Austria), as described in the manufacturer’s protocol (Molecular Probes). Equimolar amounts of the 159 PCR products from each individual were pooled to ensure equivalent hybridization. Pooled samples were purified to remove the residual primers and nucleotides using a clean-up kit (QIA Quick PCR clean-up kit; Qiagen Inc., Valencia, CA).

DNA Fragmentation, Labeling, and Hybridization to arRP-I Chips

Fragmentation of pooled samples was performed using fragmentation reagents (0.2 U DNaseI/μg DNA, from the GeneChip Resequencing Assay Kit; Affymetrix). A parallel fragmentation reaction was performed for each sample and analyzed on a 4% to 20% gradient acrylamide-TBE (Tris-boric acid-EDTA) gel followed by staining with a nucleic acid stain (SybrGold [1:50,000 in Tris-EDTA]; Molecular Probes) to evaluate the extent of fragmentation. Completely fragmented products were seen to run with 20- to 100-bp double-stranded (ds)DNA.
Sequencing Chips for Genotyping arRP Patients

Design of the Arrays

<table>
<thead>
<tr>
<th>Design of arrays</th>
<th>1.28 \times 1.28 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of an array</td>
<td>500,356</td>
</tr>
<tr>
<td>Total number of features/cells in an array</td>
<td>20 \times 25 micron</td>
</tr>
<tr>
<td>Total double stranded arRP genes sequence arrayed per chip</td>
<td>29,214</td>
</tr>
<tr>
<td>Total number of base pairs of control plasmid DNA sequenced per chip</td>
<td>25,804</td>
</tr>
<tr>
<td>Sample analysis</td>
<td></td>
</tr>
<tr>
<td>Total samples analyzed</td>
<td>35</td>
</tr>
<tr>
<td>Total base pairs of DNA sequenced on chip</td>
<td>898,650</td>
</tr>
<tr>
<td>Total arRP gene sequence assigned by GDAS</td>
<td>877,399</td>
</tr>
<tr>
<td>Total percentage bases called (%)</td>
<td>97.60%</td>
</tr>
<tr>
<td>Number of previously reported SNPs detected</td>
<td>382</td>
</tr>
<tr>
<td>Number of new polymorphisms detected</td>
<td>113</td>
</tr>
<tr>
<td>Number of previously reported mutations detected</td>
<td>4</td>
</tr>
<tr>
<td>Number of novel and potentially pathogenic changes detected</td>
<td>7</td>
</tr>
</tbody>
</table>

Replicate experiments

- **First set**
  - Total number of chips used for the analysis: 2
  - Number of known mutations tested (listed in table 4): 30
  - Number of mutations detected on chip 1: 30
  - Number of mutations detected on chip 2: 30

- **Second set**
  - Total chips analyzed for replicate experiments: 3
  - Total number of base pairs analyzed: 10,069
  - Base pairs assigned on chip 1: 9,816 (97.5%)
  - Base pairs assigned on chip 2: 9,741 (96.7%)
  - Base pairs assigned on chip 3: 9,845 (97.8%)

- **Number of sequence changes detected in this set (listed in table 4):**
  - Number of sequence changes detected on chip 1: 10
  - Number of sequence changes detected on chip 2: 10
  - Number of sequence changes detected on chip 3: 10

- **Total number of chips used for the analysis:** 2
- **Total number of mutations detected on chip 1:** 30
- **Total number of mutations detected on chip 2:** 30
- **Total number of chips used for the analysis:** 2
- **Total number of mutations detected on chip 1:** 30
- **Total number of mutations detected on chip 2:** 30

**RESULTS**

**Design of the Arrays**

The 11 genes we selected carry at least 180 previously described mutations associated with early-onset retinal degenerations (HGMD, Human Gene Mutation Database; http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html/ provided in the public domain by the Institute of Medical Genetics, University of Wales College of Medicine, Cardiff, Wales, UK). These genes were arrayed on the 30-kb resequencing platform available through Affymetrix (Table 2). The array-based sequencing depends on differential hybridization of genomic fragments to short perfect-match and mismatch oligonucleotide probes. Exons with homologous sequences cross-hybridize on the chips and consequently are not suitable for sequencing with this technology. Sequences containing repeat regions are also not suitable. Analysis of the sequence on RepeatMasker revealed that the sequence of the coding region of PDE6a and PDE6b are highly homologous, except for exon 7. Thus, we selected the sequences of all exons of the gene PDE6b, because it contains 20 known recessive RP mutations, whereas PDE6a carries only three mutations, one each in exons 7, 13, and 14. Only the nonhomologous exon 7 of the PDE6a gene was included on the chips. Exons 5 and 5 of CNGA1 and exon 5 of TULP1 contain repeat sequences and were not included.
the other exons in the 11 genes were found to be suitable for sequencing with the DNA sequencing arrays.

Fifteen-base pair flanking intron sequences on each side of each exon were included on the chips. Because the first and final 11 bp of each fragment cannot be read, we were able to analyze only 4 bp of the flanking intronic sequence from each splice junction. The total number of bases of double-stranded arRP gene sequence arrayed on the chips is 29,214. However, because 11 bp from the beginning and the end of each fragment tiled on arrays cannot be read, a total of 25,804 bp sequence of arRP genes, which includes sequence of all exons and 4-bp flanking intronic sequence, can be analyzed on each arRP-I chip. As a positive control, an 814-bp sequence from the plasmid TAG IQ-EX (Affymetrix) was tiled onto the arrays. The arRP-I chips were designed with 20 multiple copies of a specific 25-mer-oligonucleotide probe. The probes are tiled both for sense and antisense strands of the gene chip tests for the wild-type nucleotide, for substitutions with the other three bases and for a single base deletion. Therefore, a total of eight probes are tiled for each nucleotide position in addition to two ‘del’ cells corresponding to each of the four bases and a blank. These selected parts of the chips show sequence changes in the heterozygous state in four samples analyzed using the arRP-I chips.

Table 3. Summary of Novel Potentially Pathogenic Nucleotide Changes and Previously Reported Mutations Detected in Patient DNA, with the arRP-I Chips

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gene</th>
<th>Nucleotide Change</th>
<th>Amino Acid Change</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>KE727</td>
<td>RHO</td>
<td>C959A</td>
<td>Thr320Asn</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>KE1246</td>
<td>CRB1</td>
<td>G2473A</td>
<td>Glu825Lys</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>R165</td>
<td>TEULP1</td>
<td>IVS2+3, A &gt; G</td>
<td>Val567Met</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>R206</td>
<td>ABCA4</td>
<td>G1699A</td>
<td>Ser2145Phe</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>KE869</td>
<td>RGR</td>
<td>C754T</td>
<td>Arg167His</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>R353</td>
<td>MERTK</td>
<td>G500A</td>
<td>Arg167His</td>
<td>Hetero</td>
<td></td>
</tr>
<tr>
<td>R376</td>
<td>ABCA4</td>
<td>IVS232-2, A &gt; T</td>
<td>Hetero</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE385</td>
<td>RPE65</td>
<td>T963G</td>
<td>Asn321Lys</td>
<td>Hetero</td>
<td>28</td>
</tr>
<tr>
<td>KE1246</td>
<td>ABCA4</td>
<td>T3602G</td>
<td>Leu1201Arg</td>
<td>Hetero</td>
<td>39</td>
</tr>
<tr>
<td>KE1246</td>
<td>ABCA4</td>
<td>G5077A</td>
<td>Val1693Ile</td>
<td>Hetero</td>
<td>34</td>
</tr>
<tr>
<td>R376</td>
<td>ABCA4</td>
<td>C5327T</td>
<td>Pro1776Leu</td>
<td>Hetero</td>
<td>39</td>
</tr>
</tbody>
</table>

The called base sequence derived using these arrays calculated using ABACUS determine the base at each sequence position (Fig. 1). Having 10 features of eight sets of probes for a single-base sequence, makes it possible to genotype ~30 kb of DNA with more than 99% accuracy.32

Sequencing Using arRP-I Chips

Genomic DNAs from 35 unrelated patients with recessive RP were amplified using primers described in the Methods section. The amplified products were sequenced on the custom designed arRP-I chips. The total arRP gene sequence that can be read on the chip is 903,140 bp (25,804 x 35). However, in our analysis of these 35 patients, amplification products of exons of some samples corresponding to 4,490 bp were not included in the hybridization cocktail. Hence, a total of 898,650 bp of dsDNA sequence from 35 patients was analyzed on arRP-I chips. The exons not included in the hybridization mix were sequenced by dideoxy sequencing (model 3100 sequencer; ABI) to complete the analysis for each patient. In total the DNA analysis software (GeneChip; Affymetrix) assigned 877,399 base calls out of the 898,650 bp analyzed, which ranged from 96% to 98.5% of the total on different chips. The called base sequence derived using these arrays corresponded to the reference sequence tiled on the chip with more than 99% accuracy. These results were obtained when the chips were used within 7 months of manufacturing; the actual shelf-life of these chips is yet to be determined.

In the 903,140 bp of sequence analyzed from the 35 samples, we identified a total of 506 sequence changes, of which 382 corresponded to previously reported single nucleotide polymorphisms (SNPs) and four changes represented previously reported mutations (Tables 2, 3). An additional 113 changes corresponded to polymorphisms not previously reported, and the remaining seven nucleotide substitutions found were novel alterations (Table 3).
Validation of Array-Based Sequencing Data

Analysis of Different Types of Known Sequence Changes. To illustrate the application of this arRP chip approach, we used DNA from 30 selected patients comprising 30 known mutations in the genes that were tiled on the arRP-I chip (Table 4). Exons containing these mutations were selectively amplified from patients’ DNA, pooled and sequenced on two arRP-I chips. All 30 mutations were detected on both chips.

As an independent verification of the chip's ability to detect mutations, four additional DNA samples known by previous sequencing to carry 10 mutations in three genes were studied. The amplified products were pooled and analyzed on arRP-I chips in triplicate. In this trial, all 10 mutations were detected on all three chips (Table 4; Fig. 2).

As an internal control exon 25 of ABCA4 from a patient with the known mutation C3758T was substituted for exon 25 for each of the 35 patient samples. This mutation was correctly detected in all the samples. The omitted exon 25 from the patient samples was sequenced separately by the dideoxy chain termination method, to complete the analysis for each patient.

Hybridization in the Absence of a Specific Sequence in the Sample. A sample containing amplified exons of genes RPE65, RHO, and ABCA4, without the remaining eight genes, was tested on three chips. No nonspecific hybridization was observed—that is, no base calls were assigned to the sequence corresponding to the exons that were not included in the test sample (Fig. 3). These results suggest a near absence of non-specific hybridization.

Comparison with Sequence Generated Using the Dideoxy Chain Termination Method. Sequence analysis (model 3100 sequencer; ABI) was performed on 245 exons covering 45,195 bp of sequence. Among these, nucleotide changes were detected in 86 exons by the arRP-I chip. These include the sequence changes listed in Tables 2 and 3 and an additional 35 SNPs. All 86 changes were confirmed by dideoxy sequencing. No discrepancy was detected between the sequence generated using the dideoxy chain termination method and the sequence obtained with the arRP-I chips.

Array-Based Sequencing Data Output Format

Sequence output was generated in FASTA format, and the sequence of several samples were aligned on the resequencing window at the same time, along with the reference sequence. Sequence alterations are shown in IUPAC (International Union of Pure and Applied Chemistry) code and color-coded as in Figure 4. “No call” (n) was assigned to a base when the algorithm could not designate as a result of saturation of the hybridization signal, large signal-to-noise ratio, or weak signal. Analysis of the bases that were labeled n revealed that 98.4%...
were either "g" or "c", and 89.5% of these resided in stretches of G or C. In addition, if two sequence variants were located in adjacent positions, the GDAS software called the first base appropriately but the second base was often a "no call" (Fig. 4).

Comparison of hybridization signal intensities on the probe intensity window with corresponding perfect match and mismatch probes can help resolve such "no call" problems (Fig. 2).

Mutation Analysis of Samples Using the arRP-I Gene-Sequencing Arrays

We analyzed 35 patients with a diagnosis of arRP using the arRP-I sequencing arrays. Seven novel sequence alterations involving either an amino acid change or splice site were detected in seven patients, whereas four previously known mutations were detected in three (Table 3; Fig. 1). Single-nucleotide changes were detected in the heterozygous state in five patients: KE727, KE869, R353, R206, and KE385. The nucleotide change IVS2+3, A→G was detected in the homozygous state in the TULP1 gene in one patient (R165). A novel potential splice site altering heterozygous change, IVS2-2, A→T was detected in ABCA4 gene in patient R376, along with another known ABCA4 mutation in the heterozygous state. The two potential splice changes observed were tested using splice prediction programs, and the IVS2-2, A→T change in the ABCA4 gene was predicted to result in the loss of a splice acceptor site, whereas the sequence change IVS2+3 in the TULP1 gene was not predicted to alter splice sites. Patient KE1246 carried two previously reported ABCA4 mutations in the heterozygous state in addition to a novel amino acid-altering change in the CRB1 gene (Table 3).

We analyzed the available relatives of the patients listed in Table 3. Blood samples of unaffected parents and three unaffected siblings of patient KE727 were available. The mother and all three siblings carried the novel heterozygous Thr320Asn change in the RHO gene but the father did not. Samples of parents, one affected sibling and five unaffected siblings of KE869, were available. Analysis of these revealed the presence of the RGR Ser245Phe change in all six siblings in the heterozygous state and in the mother in the homozygous state. This sequence change was not detected in the DNA of the father. The mother with the RGR change in the homozygous state was reported to be unaffected at age 81 years. DNA samples from two affected siblings, two unaffected siblings, and the mother of patient R376 were obtained for the study. The two sequence changes observed in the ABCA4 gene in R376 were also detected in the two affected siblings in the heterozygous state. The change IVS2-2, A→T was detected in the heterozygous state in the mother and an unaffected sister. The other unaffected sibling, a brother, did not carry either of the changes observed in the proband. Relatives of the remaining subjects listed in Table 3 were not available for the study. Large amounts of data available through previous publications, mutation databases, an SNP database, and unpublished data from our laboratory (Downs et al., manuscript in preparation) indicate that the observed seven novel variants are rare and that some of these could be associated with pathogenicity.34,35

FIGURE 3. Part of the arRP-I chip after hybridization. DNA fragments corresponding to a few genes were not added to the hybridization reaction and no nonspecific hybridization was observed in regions corresponding to those gene sequences. Positive controls tiled on the chip for orientation (gridding) and hybridization, are shown.

FIGURE 4. Sequence output files of samples analyzed on the arRP-I sequencing arrays. Part of the sequence containing nucleotide alterations is shown. Reference sequence and positions of nucleotides are shown in red at the top. n, the intensity of the signal did not allow for a specific base call and a comparison of signal intensity is necessary to decide on the base call; y (in orange), a C or T nucleotide change in the heterozygous state; r (in orange), an A or G nucleotide change in the heterozygous state; g (in green), a nucleotide change to homozygous G.
DISCUSSION

Microarray sequencing technology is a promising new diagnostic tool. Analysis of the arRP-I chips indicates that these sequencing arrays offer a rapid, cost-effective method for genotyping patients with the genetically heterogeneous disease arRP.

The arRP-I chips we designed produced 97.6% of the sequence analyzed with >99% accuracy and reproducibility. Similar resequencing arrays for the SARS coronavirus genome and the human mitochondrial genome have been described recently. The sequence obtained using the SARS-CoV arrays was reported to be more than 99.99% accurate and reproducible, and the MitoChip produced sequence with 96% accuracy and more than 99.99% reproducibility. Our arRP-I gene chips detected sequence changes involving any combination of nucleotide exchange in the homozygous and heterozygous state. Overall performance of the custom-designed arRP-I sequencing arrays was comparable to the two previously reported custom-designed sequencing chips.

The sequencing arrays are most appropriately used for patients with phenotypes that have been previously associated with the arrayed genes. Because the genes we selected for array can cause a spectrum of disease phenotypes, one may consider screening patients beyond the previously recognized phenotypes. These chips offer a fast and efficient method of screening for mutations in most recessive RP genes. Single large genes arrayed on these chips, such as ABCA4, can also be analyzed for mutations in patients with other relevant phenotypes. Chips containing previously described mutations in the ABCA4 gene have been designed and successfully used to detect these mutations in patients with Stargardt’s macular degeneration, cone–rod dystrophy, and arRP. As observed in the patient samples analyzed using the arRP-I sequencing arrays, novel sequence alterations are likely to exist in patients with retinal conditions. In contrast to the previously reported ABCA4 chips, the arRP-I chips can detect both novel and previously known single-nucleotide changes in the sequence. Therefore, failure to detect mutations in a single gene or a group of genes using the arRP-I sequencing arrays will suggest exclusion of those genes for carrying mutations in the sequence tiled.

The arRP-I sequencing arrays offer for the first time an opportunity to screen simultaneously on a single platform for sequence alterations in multiple genes. This method of sequencing will allow us to address important questions about both simple Mendelian and complex forms of retinal degenerations involving sequence changes in more than one gene. A broad phenotypic variation within and between families with the same genotype has been described in several retinal diseases. In addition, a digenic pattern of inheritance of RP, and triallelic involvement in Bardet–Biedl syndrome have been reported. Our arRP-I chips may serve as valuable tools to evaluate the genetic basis of such complex phenotypes and to determine gene interactions.

Analysis of samples using the arRP-I sequencing arrays resulted in the identification of an average of 15 sequence changes in each sample. Some of the sequence changes detected are likely to be mutations associated with the phenotype. In addition, the disease-causing mutation may lie outside the putative coding region and the flanking regions tiled on the chips. A perceived problem of multigene sequencing is the interpretation of the results, as there is much intrinsic variation in the genome. The data generated by these high-throughput methodologies, in the long run, could be a valuable resource to study the phenotype–genotype associations. In addition, the polymorphisms detected may play a role as modifiers, either by directly influencing the function of the gene or by interacting with the environment. Careful evaluation of these sequence changes will provide an opportunity to ask questions about gene interactions and the role of these sequence alterations as modifiers of phenotype.

Generation of sequences by using the DNA array technology is simple, and the methodology is easy to establish. However, as is true of other sequencing methods, interpretation of results is time consuming. Although the data are generated in FASTA format, they do not directly interface with the sequence data available in public domains. Therefore, the interpretation of each sequence change must be performed manually. Some sequence stretches are not suitable for arrays, and these must be screened for mutations by an alternate method. Heterozygous deletions and insertions cannot be detected using resequencing arrays as well as changes in the other genomic regions (intergenic or intronic), which are not tiled on the chip. Despite these limitations, these arRP-I resequencing arrays offer an efficient and reliable method of high-throughput screening for mutations in genetically heterogeneous diseases such as arRP.

Comparison of the cost of sequencing using the arRP-I chips with the dyeoxy sequencing revealed that the material cost of sequencing by arRP-I chips is 23.5% cheaper. However, sequencing using the arRP-I chips was found to be highly cost-effective when the labor and time required for the analysis were compared. With the arRP-I arrays, all 155 exons can be analyzed in a single assay, cutting down the time required for the analysis significantly.

Currently, the size of the platforms that are available limits our ability to array all known candidate genes for the arRP phenotype on a single matrix. However, larger platforms (300-kb) are likely to be available in the near future, which will solve this problem. Sequencing arrays are likely to play a major role in genotyping and molecular characterization of patients with hereditary diseases.

Acknowledgments

The authors thank patients and their families for participating in the study; Randall Wallach, Deborah Eadie, and Regina Belaney for assistance in preparing the manuscript; and Miriam Meisler for critically reading the manuscript.

References

8. Morimura H, Fishman GA, Grover SA, Fulton AB, Berson EL, Dryja TP. Mutations in the RPDR5 gene in patients with autosomal reces-


34. Stone EM. Finding and interpreting genetic variations that are important to ophthalmologists. Trans Am Ophthalmol Soc. 2003;101:437–484.


