

Clinically Focused Molecular Investigation of 1000 Consecutive Families with Inherited Retinal Disease

Edwin M. Stone, MD, PhD,^{1,2} Jeaneen L. Andorf, BA,^{1,2} S. Scott Whitmore, PhD,^{1,2} Adam P. DeLuca, PhD,^{1,2} Joseph C. Giacalone, BS,^{1,2} Luan M. Streb, BA,^{1,2} Terry A. Braun, PhD,^{1,2} Robert F. Mullins, PhD,^{1,2} Todd E. Scheetz, PhD,^{1,2} Val C. Sheffield, MD, PhD,^{1,2,3} Budd A. Tucker, PhD^{1,2}

Purpose: To devise a comprehensive multiplatform genetic testing strategy for inherited retinal disease and to describe its performance in 1000 consecutive families seen by a single clinician.

Design: Retrospective series.

Participants: One thousand consecutive families seen by a single clinician.

Methods: The clinical records of all patients seen by a single retina specialist between January 2010 and June 2016 were reviewed, and all patients who met the clinical criteria for a diagnosis of inherited retinal disease were included in the study. Each patient was assigned to 1 of 62 diagnostic categories, and this clinical diagnosis was used to define the scope and order of the molecular investigations that were performed. The number of nucleotides evaluated in a given subject ranged from 2 to nearly 900 000.

Main Outcome Measures: Sensitivity and false genotype rate.

Results: Disease-causing genotypes were identified in 760 families (76%). These genotypes were distributed across 104 different genes. More than 75% of these 104 genes have coding sequences small enough to be packaged efficiently into an adeno-associated virus. Mutations in *ABCA4* were the most common cause of disease in this cohort (173 families), whereas mutations in 80 genes caused disease in 5 or fewer families (i.e., 0.5% or less). Disease-causing genotypes were identified in 576 of the families without next-generation sequencing (NGS). This included 23 families with mutations in the repetitive region of RPGR exon 15 that would have been missed by NGS. Whole-exome sequencing of the remaining 424 families revealed mutations in an additional 182 families, and whole-genome sequencing of 4 of the remaining 242 families revealed 2 additional genotypes that were invisible by the other methods. Performing the testing in a clinically focused tiered fashion would be 6.1% more sensitive and 17.7% less expensive and would have a significantly lower average false genotype rate than using whole-exome sequencing to assess more than 300 genes in all patients (7.1% vs. 128%; $P < 0.001$).

Conclusions: Genetic testing for inherited retinal disease is now more than 75% sensitive. A clinically directed tiered testing strategy can increase sensitivity and improve statistical significance without increasing cost. *Ophthalmology* 2017;124:1314-1331 © 2017 by the American Academy of Ophthalmology



Supplemental material available at www.aaojournal.org.

See Editorial on page 1254.

When the first inherited retinal disease genes were discovered in the late 1980s and early 1990s,¹⁻³ ophthalmic genetics largely was a descriptive subspecialty. The primary goals of the ophthalmologist were to give the patient's condition a name and to try to discern the inheritance pattern so that one could give the patient and his or her family members a reasonably accurate estimate of the risk that other family members would be affected with a similar disease. At that time, the chance that a molecular diagnosis could be accomplished for the average patient with an inherited retinal disease was less than 5%, and such tests were performed by only a few research laboratories.

The main limitations to molecular diagnosis in the early 1990s were the overall lack of knowledge of the human

genome and the relatively crude and laborious methods for investigating it. An often underappreciated positive effect of those limitations was that molecular tests in the early 1990s tended to be very focused by the clinical features of the family being studied. For example, one would not have sequenced the rhodopsin gene in a person with the clinical features of Best disease, and thus would not have been in a position to observe a rare non-disease-causing polymorphism in the rhodopsin gene and incorrectly conclude that it was disease-causing in that patient.

Many things have changed in ophthalmic genetics in the past 25 years, perhaps most notably the successful use of gene therapy for inherited retinal disease,⁴⁻⁶ the more widespread availability of preimplantation genetic testing to

reduce the recurrence of severe genetic diseases, and the introduction of clustered regularly interspaced short palindromic repeats-based genome editing,^{7–9} which, when coupled with induced pluripotent stem cells^{10,11} and in vitro retinal differentiation, have the potential to generate immunologically matched genetically corrected cells for therapeutic transplantation.^{10,11} The advent of these gene-directed interventions have increased both the value and the risks of genetic testing. For these treatments to work, one must know the disease-causing gene and in some cases the exact disease-causing mutations with complete accuracy. The diagnostic goal of the clinician is no longer just to give the clinical findings a name; rather, it is to identify the patient's disease-causing genotype with sufficient accuracy that the probability that a gene-directed intervention will help the patient is significantly greater than the possibility that it will cause harm.

Fortunately, genetic testing methods also have changed dramatically in recent years. What was once considered to be the largest scientific undertaking of mankind, the sequencing of the entire human genome^{12,13} can now be accomplished in an individual patient in just a few weeks' time for a few thousand dollars. This has led some to believe that experienced clinicians are now less necessary for the care of patients with inherited diseases and that the tests themselves are so powerful that they can provide the correct answer in almost any clinical situation regardless of the quality or quantity of accompanying clinical information. Actually, the reverse is true. As genetic tests have become larger in scope and sensitivity, the need for exceptionally detailed and accurate clinical information also has increased. This is primarily because there is a lot of normal variation in the human genome—millions of genetic differences between any 2 healthy individuals—and as a result, very broad investigations will always result in multiple plausible disease-causing findings that will need to be winnowed to 1 on clinical grounds.

We undertook this study for several related purposes: (1) to determine the current overall sensitivity of genetic testing for inherited retinal disease, (2) to determine the relative frequencies of inherited retinal disorders seen in a single North American eye clinic, (3) to determine the proportions of these diseases caused by mutations in specific genes, (4) to develop a teachable algorithm for pretest clinical diagnosis, (5) to evaluate the efficiency of a clinically driven tiered genetic testing strategy, and (6) to provide practicing ophthalmologists some insight into the complexity of next-generation sequencing (NGS) data and the obligation to apply corrections for multiple measurements to these data.

Methods

Participants

This study was approved by the institutional review board of the University of Iowa and adhered to the tenets set forth in the Declaration of Helsinki. All patients seen by a single clinician (E.M.S.) in the Retina Clinic of the Department of Ophthalmology and Visual Sciences, University of Iowa, between January 2010 and June 2016, who were judged by that clinician to have a monogenic heritable component to their eye disease and who were

60 years of age or younger when first symptomatic, were offered inclusion in the study. Those who chose to participate (more than 99% of those invited) provided written informed consent. In many cases, additional family members also were invited to participate in the study at the time of the original clinic visit, at a later visit, or by sending samples and records by mail. Patients with the following clinical diagnoses were excluded from the cohort: age-related macular degeneration, central serous retinopathy, autoimmune retinal disease, and acute zonal occult outer retinopathy.

Clinical Assessment

All probands and available family members underwent a complete eye examination including visual acuity assessment, intraocular pressure measurement, evaluation of ocular motility and pupils, slit-lamp biomicroscopy, and indirect ophthalmoscopy. Most patients also underwent Goldmann perimetry, color fundus photography, and spectral-domain optical coherence tomography. A subset of patients underwent one or more of the following: assessment of color vision, reduced intensity autofluorescence imaging,¹⁴ fluorescein angiography, or electroretinography, according to International Society for Clinical Electrophysiology of Vision standards.¹⁵

Diagnostic Classification

All available historical, clinical, electrophysiologic, and imaging data from each participant were digitized and re-reviewed by a single clinician (E.M.S.) for the purpose of placing each of the 1000 probands into an objectively defined clinical category. A patient's genotype was never used to place him into a category. Even when a clinical diagnosis seemed to be incorrect after genetic testing (such as the *de novo* rhodopsin mutation that mimicked autosomal recessive early childhood onset retinal dystrophy in a young girl; patient 442 in [Table S1](#), available at www.aaojournal.org), the original clinical diagnosis was retained. The purpose of this objective assignment was to allow us to determine how many patients of 1000 would fall into each specific clinical category and which genes were responsible for disease in each of these objectively defined categories. The names and inclusion criteria for most of the 96 resulting diagnostic categories ([Fig 1](#); [Table 1](#)) are for the most part well defined in the existing clinical literature. However, in a few cases, some empiric rules were established to define the borders between categories more clearly (see "[Results](#)"). Also, the higher-order grouping of the individual categories was somewhat nonstandard and was chosen to minimize the number of decisions or clinical tests that were needed to place a patient into a category.

Disease Genes

The published literature was reviewed to identify all genes that had been shown convincingly to cause genetic retinal disease. These 305 genes ([Table S2](#), available at www.aaojournal.org) were divided into 2 groups based on whether a gene was known to cause some combination of severe progressive loss of cognition, neuromuscular control, significantly shortened life expectancy (43 genes) or not (262 genes). The published literature also was reviewed to identify the retinal phenotypes that had been associated previously with each of these 305 genes, and these data were used to associate each gene with 1 or more of the 96 diagnostic categories shown in [Figure 1](#), [Table 1](#), and [Table S2](#) (available at www.aaojournal.org). The 43 genes associated with the more severe systemic diseases were included in the analysis only when clinical features suggestive of a debilitating systemic phenotype already were manifest.

DNA Extraction

Blood samples were obtained from all probands ($n = 1000$) and available family members ($n = 2348$), and DNA was extracted by using the manufacturer's specifications for whole-blood DNA extraction using Genra System's Autopure LS instrument (Valencia, CA).

First-Tier Genetic Testing

A preliminary mutation detection probability distribution¹⁶ was established for each of the 96 clinical categories using a combination of the published literature and the anonymized summary experience of the Carver Nonprofit Genetic Testing Laboratory at the University of Iowa. These mutation detection probability distributions were used to devise focused screens designed to detect the most common disease-causing alleles of the most common genes associated with each of the diagnostic categories. These screens each used 1 or more of the following approaches: automated Sanger sequencing with an ABI 3730xl sequencer (Carlsbad, CA), allele-specific genotyping with a Fluidigm EP1 (South San Francisco, CA) an Amplification Refractory Mutation System,¹⁷ chromosomal microarray analysis, plasmid cloning of polymerase chain reaction (PCR) products followed by Sanger sequencing, or a combination thereof. Variants were considered disease-causing if they met our previously published criteria¹⁸ for an estimate of pathogenic probability (EPP) of 2 or 3. A genotype was considered convincing, and the patient included in the calculation of the solve rate for that diagnostic category, if it consisted of a heterozygous mutation with an EPP of 2 or 3 in a gene known to cause a dominant disease, a hemizygous mutation with an EPP of 2 or 3 in a gene known to cause X-linked disease, or 2 mutations (suspected to lie on separate alleles by direct observation or statistical inference) each with an EPP of 2 or 3 in a gene known to cause recessive disease.

Cloning and Sequencing of *RPGR* Exon 15

To detect mutations in the low-complexity region of *RPGR* exon 15, Sanger sequencing of TA-cloned PCR products was performed. Patient DNA was PCR amplified and the products were gel purified and TA cloned into the pCR 2.1 TOPO Vector using the TOPO TA Cloning Kit (Invitrogen, Grand Island, NY). TA-cloned PCR products were transformed using One Shot TOP10 chemically competent cells (Invitrogen). Transformed cells subsequently were streaked and cultured on AIX plates (AIX, Aachen, Germany) for blue-white colony screening. Validated clones were picked, expanded in LB broth, purified, and Sanger sequenced on the ABI 3730xl sequencer using optimized sequencing chemistry.

Next-Generation Sequencing

Whole-exome sequencing was performed using the Agilent v5 exome kit (Santa Clara, CA) with the addition of custom xGen Lockdown probes (IDT, Coralville, IA) to target regions of the genome relevant to eye disease that are not well covered in the standard exome kit. These regions cover known noncoding mutations in *CEP290*, *USH2A*, *ABCA4*, and the long- or medium-wavelength opsin cluster, in addition to insufficiently covered coding exonic sequence in genes such as *ABCC6* (all bait sequences available on request). Whole-exome sequencing was performed using the Illumina HiSeq 2500 or 4000. Whole-genome sequencing was performed using the HiSeq X (Hudson Alpha, Huntsville, AL). Sequences were aligned to the genome using BWA.¹⁹ Single nucleotide variations and small insertions and deletions were called using GATK.²⁰ Structural variants were called using Conifer²¹ and Manta.²²

Calculation of the False Genotype Rate

Genetic variations that cause rare, high-penetrance, monogenic diseases also are rare in the population, and most genotyping pipelines, including ours, remove variants that are too common to cause the rare diseases under study. For this project, the cutoff for recessive variants was set at 0.006 (the frequency of the more common well-established disease-causing mutations in *ABCA4*), the cutoff for mitochondrial variants was 0.004 (the frequency of the most common Leber hereditary optic neuropathy variant, 11778), and the cutoff for dominant disease was set at 0.0001 (the frequency of the most common well-established mutations in *RHO*). The frequency at which one would encounter a variant at or below these thresholds in healthy people is proportional to the amount of exomic sequence analyzed and was measured directly by applying the pipeline cutoff values to the whole-exome data from the 60 000 healthy individuals collected by the Exome Aggregation Consortium (ExAC²³) and mitochondrial variants observed in 32 000 healthy individuals in MitoMap.²⁴ We defined the false genotype rate (FGR) as the frequency with which one would encounter a plausibly disease-causing recessive or dominant complete genotype when sequencing the coding regions of a specific set of genes in a healthy person. We used the ExAC data to calculate FGR values for each group of genes mapped to each specific clinical category (Fig 1). The FGR conceptually is very similar to the commonly used false discovery rate. We chose the term *false genotype rate* for the current analysis because: (1) we had recessive complete genotypes for some genes and dominant mutations for others; (2) the recessive genotypes were not observed directly, but rather were modeled from data from the ExAC database; and (3) we wanted to convey fully the associated risk of incorrect genetic test results.

Calculation of Genetic Test Costs

For each diagnostic category, a specific sequence of tests was devised based on the mutation detection probability distribution¹⁶ for that category obtained from both the published literature and the anonymized experience of the Carver Nonprofit Genetic Testing Laboratory at the University of Iowa. During development, each step in the testing sequence was optimized by subjecting it to a cost analysis (available on request). For the analysis in this article, the research cost of the currently recommended sequence of tests for each patient was calculated based on their clinical diagnosis before testing (details of the specific testing order, primer sequences, PCR conditions, etc., for any diagnostic category are available from the authors on request). The current research costs of the test components are DNA extraction and quality control genetic markers, \$40; Amplification Refractory Mutation System reaction, \$38; one set of 44 alleles assayed using the Fluidigm system, \$35; bidirectional Sanger sequencing of 1 PCR amplicon, \$20; chromosomal microarray analysis, \$500; TA cloning and bidirectional sequencing of *RPGR* exon 15 codons 762 to 1100, \$650 for males and \$975 for females; whole-exome sequencing, analysis, and confirmation, \$1200; and whole-genome sequencing, \$2450.

Results

The 1000 probands in this study were from 40 different states, the District of Columbia, and 7 foreign countries (Fig S1, available at www.aaojournal.org). Four hundred eighty-nine were female and 511 were male. The average age at entry into the study was 37.3 years (36.3 years for males and 38.5 years for females); the range

was 8 months to 88 years. Plausible disease-causing genotypes were identified in 760 of these probands, 393 males and 367 females (Table S1, available at www.aaojournal.org). The average age at entry into the study was very slightly younger for those in whom a disease-causing genotype was identified (34.9 years for males and 37.7 years for females).

The clinical classification system (Fig 1; Table 1) used in this study was devised as a means for clinicians who see adults and older children with inherited retinal diseases (1) to communicate efficiently their clinical impressions to the molecular diagnostic laboratory charged with identifying the patients' disease-causing mutations and (2) to narrow the pretest hypothesis to the smallest number of genes possible at our current level of clinical understanding. For the most part, the names used to refer to the individual clinical entities in the classification system are in common clinical use, and only the higher-order grouping of these terms is in any way unusual. This grouping was chosen to keep entities with similar genetic causes as close to one another as possible in the diagnostic tree so that if the initial screening showed negative results, a laboratory could enlarge the molecular hypothesis recursively in the most statistically efficient manner. For example, *PRPH2*-associated pattern dystrophy and *ABCA4*-associated Stargardt disease can cause almost identical clinical findings in selected patients. As a result, these categories are adjacent to each other in the classification scheme. If screening results of *ABCA4* were negative, the laboratory would screen *PRPH2* (even without a dominant family history) before moving on to screen larger genomic spaces. The most clinically homogeneous and genetically heterogeneous groups were those affected with nonsyndromic acquired photoreceptor degeneration (retinitis pigmentosa, group IA1a; and cone or cone-rod dystrophy, group IA1b; Fig 1). Multiplex kindreds belonging to these large categories were subdivided according to their pedigree structure as follows: (1) X-linked (affected males in multiple sibships connected to each other through unaffected or mildly affected females with no instances of male-to-male transmission), (2) autosomal dominant (a minimum of 3 generations with at least 1 instance of male-to-male transmission), (3) autosomal recessive (multiple affected individuals in a single sibship with healthy parents), and (4) other multiplex (all other multiplex kindreds).

Placement into 1 of the first 3 categories of congenital or stationary photoreceptor disease (IA2a–c) required clear historical evidence of parental or physician awareness of significant visual dysfunction—more than just night blindness—before the patient's fourth birthday. These patients were further divided into (1) Leber congenital amaurosis if their visual acuity was so poor that they did not use it for education or activities of daily living; (2) severe early childhood-onset retinal dystrophy²⁵ if they had useful vision but became legally blind before 10 years of age; or (3) early childhood-onset retinal dystrophy if they were not legally blind before 10 years of age. Patients were diagnosed with congenital stationary synaptic dysfunction (IA2g) if they had stable reduced acuity from birth, selective loss of the B-wave on the scotopic electroretinogram, and diminished B-wave amplitudes on the photopic electroretinogram, but no difficulties with vision in dim light. As the clinical records for the 1000 patients and their relatives were reviewed to place them into these categories, it became evident that 10 types of easily obtainable historical information were of particular value in reproducibly assigning patients to these categories, and a form was created to assist physicians in acquiring these historical data in a prospective manner (Table S3, available at www.aaojournal.org).

For this study, patients could be assigned to a higher-order point in the classification system if there were insufficient data to make a more specific assignment. For example, isolated patients with

retinitis pigmentosa were assigned to IA1a, whereas a member of an autosomal dominant family with affected individuals in 3 generations and clear male-to-male transmission was assigned to IA1aii. Of the 96 possible locations a proband could be placed in this classification system, only 62 were used at least once when subdividing the 1000 probands in this study (Fig 1; Table 1).

Figure 2 shows the breakdown of the 1000 probands among the most common diagnostic categories, whereas Figure 1 shows the frequency with which a convincing disease-causing genotype could be identified in each of these categories. In this cohort, 64.7% of the probands had photoreceptor disease (category I), 28.2% had a macular dystrophy (category II), and 7.1% had 1 of the 42 entities of the category III. Overall, convincing genotypes were identified in 76% of the probands with the highest positivity among those with a macular dystrophy (88.3%). With 4 exceptions, patients with autosomal recessive disease were considered positive only if they had both disease alleles identified. The exceptions were patients with a clinical diagnosis of Stargardt disease (23 probands), Usher syndrome (2 type-1 and 5 type-2 probands), achromatopsia (1 proband), and homocystinuria (1 proband), each of whom were found to have a convincing disease-causing mutation on only 1 allele (the FGR was less than 1% for each of these pretest hypotheses).

Figure 1 also shows the genetic heterogeneity of each of the 62 clinical categories with at least 1 patient assigned to it in this study. The most heterogeneous category was simplex retinitis pigmentosa (IA1a), which had disease-causing mutations identified in 36 different genes and a total solve rate of only 56.7%. The least heterogeneous category with at least 10 probands in it was choroideremia (IIIA1), which had a 100% solve rate in 14 patients, all with mutations in a single gene (*CHM*).

Figure 3, Table 2, and Table S4 (available at www.aaojournal.org) show the frequencies of disease-causing genotypes in each of the 104 genes that were found to cause disease in at least 1 family in the cohort. *ABCA4* was the single most common disease-causing gene and was responsible for disease in 173 families. Twelve additional genes, *USH2A*, *RPGR*, *RHO*, *PRPH2*, *BEST1*, *CRB1*, *BBS1*, *CEP290*, *PRPF31*, *CHM*, *RS1*, and *RPI* each caused disease in 1% or more of the cohort, and these 13 genes collectively were responsible for disease in almost one half of the families (n = 497). The remaining 91 genes each caused disease in less than 1% of the cohort and collectively caused disease in 26.2% of the total. Thirty genes each caused disease in a single family, and 1 family had a de novo chromosomal translocation. This cohort is certainly not a random sample of the United States population. However, it was ascertained consecutively, and was drawn from 40 of the 50 states of the United States. Thus, we believed that it would be reasonable to use these data to provide a rough estimate of the total number of individuals affected with each gene-specific disease in the country. Assuming that mutations in *ABCA4* cause disease in 1 of 10000 people,²⁶ Table 2 gives an estimate for the total number of people of all ages in the United States with mutations in this gene. Similarly, by using 2010 United States census data that show 20.2 million people in the United States younger than 5 years,²⁷ one can also estimate the number of new cases of each gene-specific retinal disease in the United States per year. Collectively, these data suggest that there are currently approximately 140000 people in the United States affected with 1 of the diseases evaluated in this study and approximately 1700 new cases per year.

Figure 4 depicts a cost and yield comparison of 2 different strategies one could use for genotyping the 1000 probands of this study. In 1 case, whole-exome sequencing would be performed on every proband and the resulting data would be evaluated for mutations in the 301 nonmitochondrial genes selected for

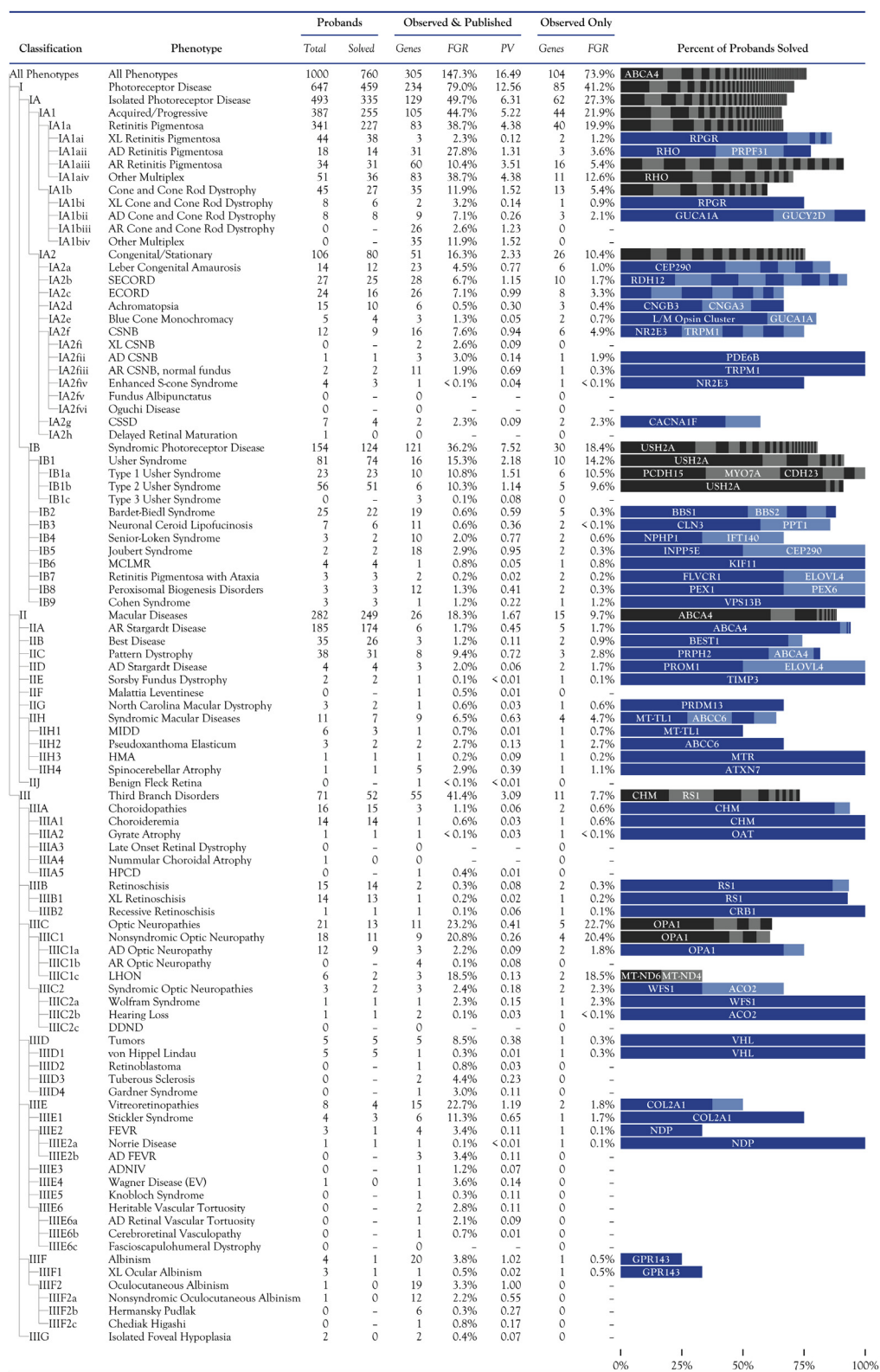


Figure 1. Table showing the distribution of patients and molecular findings across all levels of the clinical classification system. The structure of the classification system is shown at left with the common clinical terms for each phenotypic group shown in the adjacent column. The “Total” column provides the number of probands assigned to each clinical group, whereas the “Solved” column shows the number of probands in each group with a disease-causing genotype identified. The “Genes” columns provide the number of genes that have been observed to cause the diseases of that clinical group in the published literature, at the University of Iowa, or both. The false genotype rate (“FGR”) columns give the percentage of healthy individuals who would be expected to

inclusion in this study (see “Methods”). In this case, the cost to genotype each patient would be the same (\$1200; see “Methods”), and the overall yield would be 70%. The mutations that would be missed would be those that lie in noncoding sequences (e.g., nonexonic mutations that cause Stargardt disease,²⁸ Usher syndrome,²⁹ retinitis pigmentosa,³⁰ Leber congenital amaurosis,^{31,32} and North Carolina macular dystrophy³³), mitochondrial DNA (e.g., mutations that cause Leber hereditary optic neuropathy³⁴ and maternally inherited diabetes and deafness³⁵), and repetitive regions (e.g., the repetitive region of *RPGR* exon 15^{36–38}). In the other case, one would use a tiered testing strategy in which the testing was customized based on the clinical findings and testing would be stopped as soon as a complete genotype was identified. With the latter strategy, the cost would range from \$80 per patient (for those in whom a complete genotype was identified in the initial tier of testing) to more than \$2500 per patient for those who did not have a complete genotype found on prescreening and were judged suitable for whole-genome testing (see “Methods”). With this tiered strategy, the average cost per patient would be 17.7% less (\$990) than performing whole-exome sequencing in everyone and the sensitivity would be a 6.1% higher because of the findings one would make in the noncoding regions, mitochondrial genes, and repetitive DNA that were included specifically in the clinically focused tests.

Figure 5 depicts the more important difference between the 2 screening strategies: the effect on the FGR. Genetic variants that are rare enough in the general population to cause a Mendelian retinal disease are surprisingly common in whole-exome sequencing data. In this study, the population frequency cutoff was set according to the most common well-established retinal disease-causing mutations (see “Methods”). If one applies these criteria to the sequence data from the 60 000 healthy individuals in the ExAC database,²³ one observes an average of 1.28 plausible disease-causing genotypes per person among the coding sequences of the 301 nonmitochondrial candidate genes considered in this study. Another way to state this is that with a coding sequence hypothesis comprising 301 genes, there is an average FGR (see “Methods”) of 128%. For most medical tests, one would want positive results to occur by chance no more than 5% of the time, and for tests that would be used as the basis of preimplantation genetic testing or subretinal gene therapy, one might argue that it should be even less frequent.

Figure 5 shows that one can reduce the FGR to clinically useful levels by narrowing the pretest hypothesis to a relatively small number of genes. A tiered testing strategy linked to the clinical classification system in this study would identify plausible disease-causing genotypes in 48.7% of the cohort with an FGR less than 5%. Figure S2 (available at www.aaajournal.org) shows that one also can reduce the FGR per category at a given institution by first considering the genes that have been observed previously to cause disease in patients seen at that institution and then, if the results are negative, considering a larger literature-

based group of candidates and adding a statistical penalty for the additional hypothesis. The rationale for this 2-step analysis is that the previous 1000 patients seen in a given institution are likely to be more genetically similar to the next 1000 patients seen there than they will be to the entire world population represented in the published literature.

For patients whose FGR is more than 5%, which using the tiered strategy is most commonly the result of our current inability to reduce the genetic heterogeneity of categories like simplex retinitis pigmentosa on clinical grounds (Fig 1), it is especially important to confirm the phase or segregation, or both, of their putative disease-causing variant(s) and to be a bit more skeptical of molecularly weaker genotypes such as those entirely comprising novel missense variants. Table 3 shows the distribution of the 760 disease-causing genotypes identified in this study among inheritance patterns and mutation types. Of the genotypes, 2.5% involved molecularly confirmed de novo variants, which is a considerable underestimate of the actual de novo rate given that sufficient family samples to evaluate both parental alleles were available in less than 65% of families.

Five Illustrative Patients

Patient A was a 47-year-old man who first noted in his 20s difficulty following the flight of a ball after it was thrown in the air (patient 375 in Table S1, available at www.aaajournal.org). In his early 40s, he was examined and believed to have cone dystrophy. He has no family history of a similar disorder. Our examination revealed best-corrected visual acuity of 20/50 in the right eye and 20/80 in the left eye. Ophthalmoscopy revealed an iridescent golden sheen to the entire posterior pole with the exception of a reddish atrophic circular area 1.5 mm in diameter centered on the fovea in both eyes (Fig 6A). Optical coherence tomography showed a sharply demarcated loss of photoreceptors and retinal pigment epithelium (RPE) corresponding to the atrophic area seen on fundus examination (Fig 6B). Goldmann perimetry revealed a loss of the I2e isopter as well as a central scotoma to the III4e (Fig 6C). Plasmid cloning and DNA sequencing of the repetitive portion of *RPGR* exon 15 revealed a 2-base pair deletion in codon 1059. *RPGR* codons 800 through 1070 are poorly covered by whole-exome sequencing, and this mutation is undetectable with this method. It is also interesting that some frameshifting mutations in this exon are associated with a late-onset cone-selective disease,^{37,38} as seen in this patient, whereas similar mutations elsewhere in the gene cause severe rod predominant retinitis pigmentosa.

Patient B was an 8-year-old boy who first noted difficulties seeing in dim light in early childhood. His maternal grandfather had been diagnosed with choroideremia. On examination, his visual acuity was 20/32⁻¹ in the right eye and 20/32⁻² in the left eye. Ophthalmoscopy revealed extensive nummular areas of RPE and choriocapillaris loss each surrounded by a thin rim of

harbor a plausible disease-causing complete genotype by chance in any of the genes assigned to each clinical category in the published literature, at the University of Iowa, or both. “PV” is the average number of plausible disease-causing variants one would expect to observe in a healthy individual by chance in any of the genes assigned to each clinical category in the published literature. The bar lengths represent the percent of solved cases for each clinical category, whereas the alternating shades represent the proportional contributions of each gene in descending order. Gene names are given for any genes that cause at least 15% of the disease in a given category. Blue bars indicate categories with an FGR less than 5%, whereas grey bars indicate categories with an FGR of 5% or more. AD = autosomal dominant; ADNIV = autosomal dominant neovascular inflammatory vitreoretinopathy; AR = autosomal recessive; CSNB = congenital stationary night blindness; CSSD = congenital stationary synaptic dysfunction; DDND = developmental delay and/or neuromuscular degeneration; ECORD = early childhood-onset retinal dystrophy; EV = erosive vitreoretinopathy; FEVR = familial exudative vitreoretinopathy; HMA = homocystinuria with macular atrophy; HPCD = helicoid peripapillary chorioretinal degeneration; LHON = Leber hereditary optic neuropathy; MCLMR = microcephaly with or without chorioretinopathy, lymphedema and mental retardation; MIDD = maternally inherited diabetes and deafness; SECORD = severe early childhood-onset retinal dystrophy; XL = X-linked.

Table 1. Inherited Retinal Disease Categories

I – Photoreceptor disease
A – Isolated
1 – Acquired/progressive
a – Retinitis pigmentosa
i – X-linked
ii – Autosomal dominant
iii – Autosomal recessive
iv – Other multiplex
b – Cone and cone rod dystrophy
i – X-linked
ii – Autosomal dominant
iii – Autosomal recessive
iv – Other multiplex
2 – Congenital/stationary
a – Leber congenital amaurosis
b – Severe early childhood-onset retinal dystrophy
c – Early childhood-onset retinal dystrophy
d – Achromatopsia (congenital stationary cone dysfunction)
e – Blue cone monochromacy
f – Congenital stationary night blindness
i – X-linked
ii – Autosomal dominant
iii – Autosomal recessive with normal fundus
iv – Enhanced S-cone syndrome
v – Fundus albipunctatus
vi – Oguchi disease
g – Congenital stationary synaptic dysfunction
h – Delayed retinal maturation
B – Syndromic
1 – Usher syndrome
a – Type I
b – Type II
c – Type III
2 – Bardet-Biedl syndrome
3 – Neuronal ceroid lipofuscinosis
4 – Senior-Loken syndrome
5 – Joubert syndrome
6 – Microcephaly with or without chorioretinopathy
lymphedema and mental retardation
7 – Retinitis pigmentosa with ataxia
8 – Peroxisomal biogenesis disorders
9 – Cohen syndrome
II – Macular dystrophies
A – Autosomal recessive Stargardt disease
B – Best disease
C – Pattern dystrophy
D – Autosomal dominant Stargardt disease
E – Sorsby fundus dystrophy
F – Malattia leventinese
G – North Carolina macular dystrophy
H – Syndromic macular diseases
1 – Maternally inherited diabetes and deafness
2 – Pseudoxanthoma elasticum
3 – Homocystinuria with macular atrophy
4 – Spinocerebellar atrophy
J – Benign fleck retina
III – Third branch disorders
A – Choroidopathies
1 – Choroideremia
2 – Gyrate atrophy
3 – Late-onset retinal dystrophy
4 – Nummular choroidal atrophy
5 – Helicoid peripapillary chorioretinal degeneration
B – Retinoschisis
1 – X-linked
2 – Recessive
C – Optic neuropathies
1 – Nonsyndromic
a – Autosomal dominant

Table 1. (Continued)

b – Autosomal recessive
c – Leber hereditary optic neuropathy
2 – Syndromic
a – Wolfram syndrome
b – Hearing loss
D – Tumors
1 – von Hippel-Lindau
2 – Retinoblastoma
3 – Tuberous sclerosis
4 – Gardner syndrome
E – Vitreoretinopathies
1 – Stickler syndrome
2 – Familial exudative vitreoretinopathy
a – Norrie disease
b – Autosomal dominant
3 – AD neovascular inflammatory vitreoretinopathy
4 – Wagner disease (erosive vitreoretinopathy)
5 – Knobloch syndrome
6 – Heritable vascular tortuosity
a – Autosomal dominant retinal vascular tortuosity
b – Cerebroretinal vasculopathy
c – Fascioscapulohumeral dystrophy
F – Albinism
1 – X-linked ocular albinism
2 – Oculocutaneous albinism
a – Nonsyndromic
b – Hermansky-Pudlak
c – Chediak-Higashi
G – Isolated foveal hypoplasia

hyperpigmented RPE (Fig 7A). The retinal arterioles were near normal in caliber. The Goldmann visual fields were surprisingly well preserved for this degree of retinal loss (Fig 7B). His mother (patient 938 in Table S1, available at www.aaojournal.org) and sister both exhibited so-called mud-spattered pigment mottling of the fundus consistent with the carrier state of an X-linked disease. Conventional DNA sequencing failed to detect a mutation in the *CHM* gene. However, the phenotype and history were so convincing that whole-genome sequencing was performed in the child and revealed a complete duplication of *CHM* exons 6 through 8, which had been invisible to the nonquantitative PCR-based DNA sequencing.

Patient C was a 48-year-old woman whose macular pigment mottling first was noted incidentally on fundus examination at 33 years of age (patient 920 in Table S1, available at www.aaojournal.org). Her best-corrected visual acuity on examination was 20/20 in the right eye and 20/60⁺² in the left eye. Ophthalmoscopy revealed patchy loss of the RPE and choriocapillaris in the left eye more than the right eye (Fig 8A and B). Fundus autofluorescence revealed more extensive involvement than was visible ophthalmoscopically (Fig 8C and D). Gestational diabetes developed at 27 years of age, and she was diagnosed with type 2 diabetes at 32 years of age. Her mother and maternal aunt are both diabetic as well. She demonstrated hearing loss in her mid 30s and now wears hearing aids. Polymerase chain reaction–based conventional DNA sequencing revealed a heteroplasmic mutation in the mitochondrial DNA at position 3243, which is known to cause an atrophic maculopathy with maternally inherited diabetes and deafness.³⁵ Whole-exome sequencing does not assess mitochondrial DNA routinely, and as a result, this mutation would have been missed unless it was specifically sought because of her phenotype.

Patient D was a 10-year-old girl who first had difficulty seeing the blackboard in school at 7 years of age. She had a family history of a similar disease in her father. Her best-corrected visual acuity

(Continued)

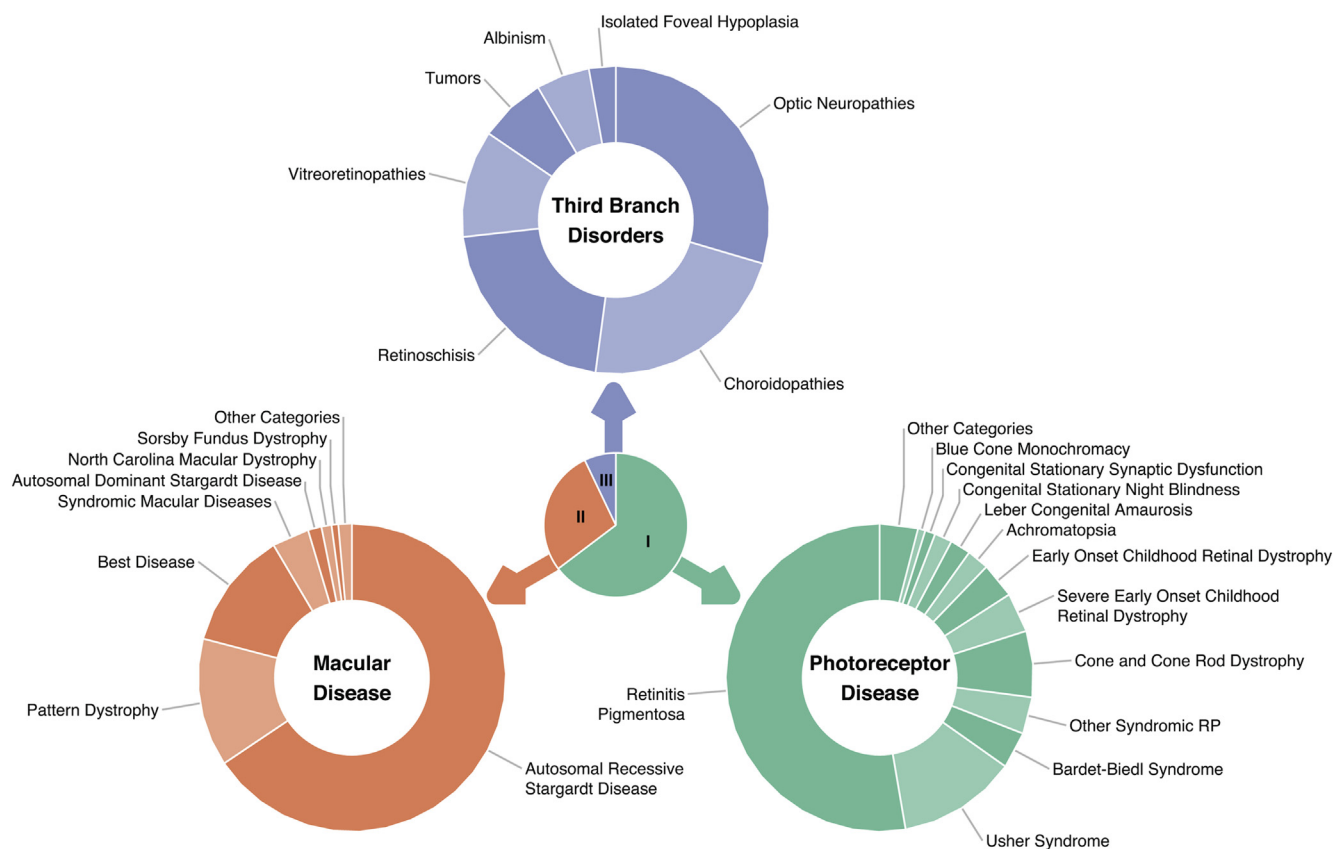


Figure 2. Graphical depiction of the distribution of 1000 consecutive probands among the larger diagnostic categories. The center chart indicates the proportion of probands assigned to each of the 3 main branches of the classification system. The outer charts show the fraction of probands assigned to the larger diagnostic categories within each branch. RP = retinitis pigmentosa.

on our examination was 20/200⁺² in the right eye and 20/200⁺¹ in the left eye. Ophthalmoscopy revealed a circular area of RPE atrophy 1 mm in diameter centered on fixation in both eyes and yellow pisciform flecks throughout the posterior pole in both eyes (Fig 9A). Optical coherence tomography revealed thinning of the outer nuclear layer and disruption of the ellipsoid zone in an area somewhat larger than the area of RPE atrophy (Fig 9B). Goldmann perimetry results were normal except for small central scotomas to the I4e in both eyes (Fig 9C). Sanger sequencing of the coding portions of the *ABCA4* gene revealed a single heterozygous missense mutation (Leu2229Pro). Sequencing of nonexomic regions previously shown to harbor disease-causing mutations revealed a previously described²⁸ cryptic splice activator on the allele opposite the missense variation (IVS36+1216 C→A). The nonexomic mutation would not have been captured by any currently available commercial exome capture reagents.

Patient E was the 42-year-old father of patient D who first noted difficulty with his central vision at 6 years of age (patient 804 in Table S1, available at www.aaojournal.org). The following year he was diagnosed with Stargardt disease, and by 13 years of age, his acuity had fallen to 20/400 in the right eye and 20/240 in the left eye. On this visit, his acuity was 20/800 in the right eye and 20/250 in the left eye. Ophthalmoscopy revealed an elliptical zone of RPE and choriocapillaris atrophy centered on fixation, very narrowed arterioles, and extensive bone-spicule-like pigment in the midperiphery in both eyes (Fig 10A). Optical coherence tomography (Fig 10B) revealed preservation of inner retinal lamination even in the area of macular atrophy. Goldman

perimetry revealed complete loss of sensitivity to the I4e stimulus throughout the visual field and an absolute scotoma inferonasally in both eyes (Fig 10C). This patient shared the IVS36+1216 C→A nonexomic mutation with his daughter and harbored an Arg2077Trp variant on his other allele. Schindler et al³⁹ found the Arg2077Trp variant to be the most severe Stargardt allele of the 16 they evaluated. This is consistent with the more severe retinitis pigmentosa-like phenotype in this individual.

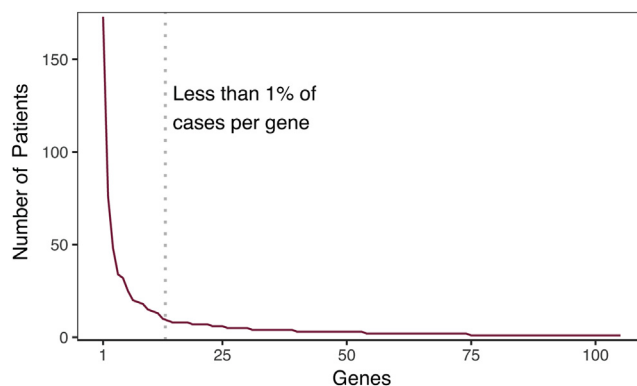


Figure 3. Graph showing the distribution of the number of probands per gene. Thirteen genes each caused disease in 1% or more of the probands in this study (left of dashed vertical line), whereas the other 91 each caused disease in less than 1%. These data are presented in more detail in Table 2.

Table 2. Estimate of the Total Number of People of All Ages in the United States with Mutations in Genes Observed in This Study

Gene	No. in Cohort	Frequency in the United States	No. in the United States	New Cases per Year
ABCA4	173	1/10 000	32 440	404
USH2A	76	1/22 763	14 251	177
RPGR	48	1/36 042	9001	112
RHO	34	1/50 882	6376	79
PRPH2	32	1/54 062	6000	75
BEST1	25	1/69 200	4688	58
CRB1	20	1/86 500	3750	47
BBS1	19	1/91 053	3563	44
CEP290	18	1/96 111	3375	42
PRPF31	15	1/115 333	2813	35
CHM	14	1/123 571	2625	33
RS1	13	1/133 077	2438	30
RP1	10	1/173 000	1875	23
FAM161A	9	1/192 222	1688	21
MYO7A	8	1/216 250	1500	19
OPA1	8	1/216 250	1500	19
PCDH15	8	1/216 250	1500	19
RP2	8	1/216 250	1500	19
GUCA1A	7	1/247 143	1313	16
IMPG2	7	1/247 143	1313	16
MAK	7	1/247 143	1313	16
PDE6B	7	1/247 143	1313	16
EYS	6	1/288 333	1125	14
PROM1	6	1/288 333	1125	14
RDH12	6	1/288 333	1125	14
CLN3	5	1/346 000	938	12
CNGB3	5	1/346 000	938	12
IQCB1	5	1/346 000	938	12
NR2E3	5	1/346 000	938	12
VHL	5	1/346 000	938	12
BBS2	4	1/432 500	750	9
CACNA1F	4	1/432 500	750	9
CDH23	4	1/432 500	750	9
CDHR1	4	1/432 500	750	9
FLVCR1	4	1/432 500	750	9
GUCY2D	4	1/432 500	750	9
KIF11	4	1/432 500	750	9
KLHL7	4	1/432 500	750	9
NMNAT1	4	1/432 500	750	9
BBS10	3	1/576 667	563	7
CERKL	3	1/576 667	563	7
CNGA3	3	1/576 667	563	7
COL2A1	3	1/576 667	563	7
CRX	3	1/576 667	563	7
ELOVL4	3	1/576 667	563	7
IFT140	3	1/576 667	563	7
INPP5E	3	1/576 667	563	7
L/M Opsin Cluster	3	1/576 667	563	7
MERTK	3	1/576 667	563	7
MT-TL1	3	1/576 667	563	7
PRPF8	3	1/576 667	563	7
RPE65	3	1/576 667	563	7
VPS13B	3	1/576 667	563	7
ABCC6	2	1/865 000	375	5
ACO2	2	1/865 000	375	5
ADGRV1	2	1/865 000	375	5
CNGB1	2	1/865 000	375	5
DHDDS	2	1/865 000	375	5
IMPDH1	2	1/865 000	375	5
KCNV2	2	1/865 000	375	5
MKKS	2	1/865 000	375	5
NYX	2	1/865 000	375	5
PEX1	2	1/865 000	375	5

Table 2. (Continued.)

Gene	No. in Cohort	Frequency in the United States	No. in the United States	New Cases per Year
PPT1	2	1/865 000	375	5
PRDM13	2	1/865 000	375	5
PRPF3	2	1/865 000	375	5
RPGRIPI	2	1/865 000	375	5
SNRNP200	2	1/865 000	375	5
TIMP3	2	1/865 000	375	5
TRNT1	2	1/865 000	375	5
TRPM1	2	1/865 000	375	5
USH1C	2	1/865 000	375	5
WDR19	2	1/865 000	375	5
ZNF408	2	1/865 000	375	5
ABHD12	1	1/1 730 000	188	2
AIPL1	1	1/1 730 000	188	2
ATXN7	1	1/1 730 000	188	2
BBS9	1	1/1 730 000	188	2
CABP4	1	1/1 730 000	188	2
CEP78	1	1/1 730 000	188	2
CLRN1	1	1/1 730 000	188	2
GPR143	1	1/1 730 000	188	2
HADHA	1	1/1 730 000	188	2
IFT172	1	1/1 730 000	188	2
Karyotypic	1	1/1 730 000	188	2
LCA5	1	1/1 730 000	188	2
MAN2B1	1	1/1 730 000	188	2
MFRP	1	1/1 730 000	188	2
MFSN8	1	1/1 730 000	188	2
MT-ND4	1	1/1 730 000	188	2
MT-ND6	1	1/1 730 000	188	2
MTR	1	1/1 730 000	188	2
NDP	1	1/1 730 000	188	2
NPHP1	1	1/1 730 000	188	2
OAT	1	1/1 730 000	188	2
PAX6	1	1/1 730 000	188	2
PEX6	1	1/1 730 000	188	2
PNPLA6	1	1/1 730 000	188	2
POMGNT1	1	1/1 730 000	188	2
RLBP1	1	1/1 730 000	188	2
RPIA	1	1/1 730 000	188	2
SLC24A1	1	1/1 730 000	188	2
TULP1	1	1/1 730 000	188	2
USH1G	1	1/1 730 000	188	2
WFS1	1	1/1 730 000	188	2

Discussion

Data that are used to arrive at a diagnosis are often incomplete, noisy, and somewhat biased. After a diagnosis is made, treatment outcomes also depend on individual patient variation, the point in the disease course that a treatment is administered, and in some cases, the skill of a surgeon in delivering a treatment to the desired anatomic location. Most physicians effectively combat these challenges with systematic actions, good record keeping, and periodic review of their outcomes in the context of new knowledge. The purpose of this study was to review the clinical and molecular findings from 1000 consecutive families affected with inherited retinal disease—in the context of current technology, public databases, and literature—to identify opportunities for improving our accuracy and efficiency in arriving at clinical and molecular diagnoses for patients with inherited retinal diseases. The consecutive nature of the ascertainment allows a rough

approximation of the total numbers of individuals in the United States who are affected by various categories of disease (Table 2). These data may be useful as scientists try to devise and implement practical comprehensive strategies for reaching all such patients with some type of useful treatment.

The clinical classification system used in this study (Fig 1; Table 1) is an empiric, internally consistent shorthand that can be used to communicate clinical observations to the laboratory efficiently for the purpose of guiding their molecular investigations, analyses, and interpretations and to align the resulting genotype–phenotype correlations with the constantly changing medical literature. This system was devised by a single clinician over many years and should not be considered a consensus view of how these disorders can be arranged most meaningfully. It is expected and desirable that other physicians will add or subtract categories from this classification scheme as needed to encompass the patients they see in their practice and to

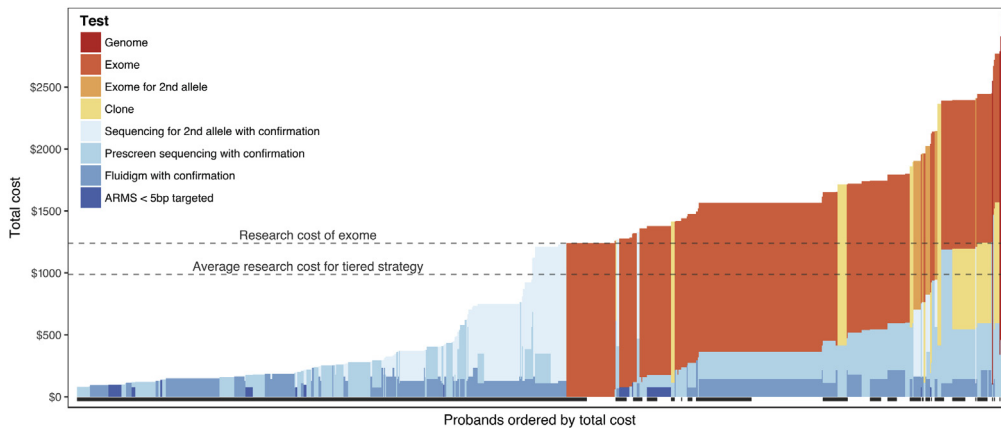


Figure 4. Graph showing the financial cost and diagnostic yield of a tiered testing strategy. Patients are ordered from lowest cost to highest cost, with colors representing the component costs of our currently recommended series of genetic tests for each clinical category. A black bar beneath a patient indicates that a causative genotype was discovered in this individual. The horizontal lines highlight the higher cost of uniform whole-exome sequencing (upper line) as compared with the average cost of clinically focused individualized tests (lower line). ARMS = amplification refractory mutation system.

move the clinical entities around to reflect better the order in which they typically pursue a diagnostic workup and the specific diagnostic instruments routinely available to them. The power of this approach lies not in the details of the classification system, but rather in the idea of using clinical information to narrow the pretest hypothesis for the purpose of increasing the sensitivity of the testing and dramatically increasing the statistical significance of the results. To reduce the FGR to less than 5%, which would be desirable when contemplating a significant intervention such as gene replacement therapy or the preimplantation selection of embryos for disease avoidance, one would need to reduce the pretest hypothesis in most cases to fewer than 10 genes (e.g., category IA2b in Fig 1). More than 85% of the terminal categories in the current classification scheme have

an FGR of 5% or less (blue bars in Fig 1). The remaining task for clinicians who care for patients with inherited retinal diseases is to scrutinize carefully the ones in the more genetically heterogeneous categories (grey bars in Fig 1) for subtle clinical signs that can be used to subdivide them further into entities associated with a smaller number of genes. Over time, some diagnostic categories and classification arrangements will prove more useful than others for this purpose, and an optimal scheme for all inherited eye diseases can evolve by combining the best features of many classifications based on their performance in the pretest prediction of the patients' genotypes.

There are many different strategies that one can use to analyze a patient's DNA for the presence of disease-causing sequence variations, and a complete discussion of them is

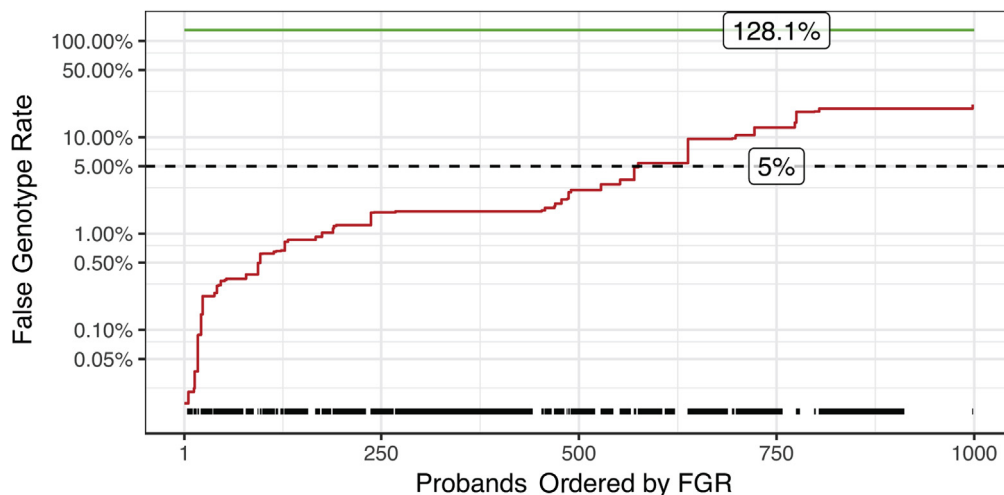


Figure 5. Graph showing statistical costs. The false genotype rate (FGR) is the average number of complete genotypes one would expect to observe by chance in a healthy individual in a specified genomic space, based on data from 60 000 healthy individuals.²³ The probands in this study are shown ordered according to the FGR associated with their clinical category (see Fig 1). The red line indicates the FGR associated with the genes observed to cause disease in this cohort (see also Figure S2, available at www.aaojournal.org). The dashed line indicates an FGR of 5% (i.e., the threshold at which 1 in 20 people would be expected to harbor a plausibly pathogenic, complete genotype by chance). The black bars at the bottom of the figure indicate that a disease-causing genotype was identified in this proband. Assessing the coding sequences of all 301 nonmitochondrial genes in all probands (green line) would result in an average FGR of 128% (i.e., these probands would be expected to harbor an average of 1.28 plausible, complete genotypes by chance).

Table 3. Variation Distribution across Inheritance Types

Inheritance Type	Missense	Terminating	Missense/Missense	Missense/Terminating	Terminating/Terminating	Totals
Autosomal dominant	140	32	—	—	—	172
X-linked	20	74	—	—	—	94
Mitochondrial	5	—	—	—	—	5
Autosomal recessive 1-allele identified	17	15	—	—	—	32
Autosomal recessive	—	—	146	160	151	457
Totals	182	121	146	160	151	760

— = not applicable.

well beyond the scope of this article. For the present purpose, it is sufficient to think of the many possibilities in terms of 4 attributes: (1) the degree to which a test can be customized to detect specific variations that otherwise would be missed, (2) the degree to which the test yields a dataset that can be reanalyzed in the future to discover currently unrecognized pathogenic variations, (3) the degree to which multiple platforms are used to maximize the strengths and minimize the weaknesses of each, and (4) the degree to which the patients' true disease-causing genotypes will be obscured by normal, non-disease-causing genetic variation.

Next-generation sequencing panels have been designed for many diseases and have the advantages that they are relatively focused (compared with whole-exome or whole-genome tests), they can be customized to include specific nonexomic regions known to cause disease, and they are relatively quick and inexpensive to perform. The disadvantage of such panels is that when showing negative results, they do not allow wider analytical exploration in search of disease-causing mutations outside the genomic space covered by the panel's design. These panels have difficulty in accurately detecting variants within repetitive DNA sequences and can have difficulty detecting deletions larger than 100 nucleotides and smaller than a few exons. Moreover, most of these panels evaluate a sufficient number of genes such that the FGR associated with them is more than 5% unless the ordering physician controls this by making a firm and narrow pretest diagnosis and rigorously evaluates the results in that context.

Whole-exome sequencing has the advantage of sampling nearly all of the transcribed sequences in the human genome and can be subjected to very focused analysis to yield statistically meaningful results. If such a focused analysis shows negative results, the data can be reanalyzed to consider a larger portion of the exome, reanalyzed at a later date when new regions of the exome may have been discovered to cause a phenotype similar to the patient under study, or both. The disadvantages of whole-exome sequencing are that it is more expensive and time consuming to perform than an NGS panel and most commercial whole exome reagents are not easily customizable to analyze specific nonexomic regions of interest to specific subspecialties of medicine. Whole-exome sequencing also has difficulty with repetitive DNA and can have even greater difficulty detecting single-exon deletions than NGS-based panels.^{40,41} As shown in Figure 5, unless one establishes a

narrow pretest hypothesis and evaluates the results accordingly, whole-exome sequencing frequently will have an FGR that is so high that the results should be considered hypothesis generating at best.

Whole-genome sequencing evaluates nearly all of the nonrepetitive sequences in the genome, and although it examines more than 50 times more of the sequence than whole-exome sequencing, is surprisingly only approximately twice as expensive as the latter method. It is better at detecting deletions, duplications, and inversions than whole-exome sequencing⁴² and also can detect disease-causing variations in nonexomic space.^{33,43–45} However, the amount of background genetic variation in the nonexomic space is so large, and our current understanding of the function of nonexomic sequences currently is so limited, that pathogenic single nucleotide variations will be completely hidden in the noise unless the pretest hypothesis is limited to only 1 or 2 genes and some functional test can be used to validate the findings.^{29,33,46} For example, the identification of a number of nonexomic mutations in *ABCA4*²⁸ required access to a large cohort of patients with convincing clinical characteristics of Stargardt disease and only a single disease-causing mutation, as well as a rather narrow mechanistic hypothesis, altered splicing, coupled with a convincing assay of this mechanism. Similarly, the discovery of the nonexomic mutations responsible for North Carolina macular dystrophy required decades of clinical and molecular genetic work to narrow the genetic interval to less than 1 million base pairs as well as sufficient families to identify 3 different mutations tightly clustered in a single regulatory element.³³

It is also important to remember that none of the commonly used high-throughput sequencing methods can distinguish unambiguously whether 2 different mutations observed in a patient were inherited from a single parent, which would not be expected to cause autosomal recessive disease, or whether they were inherited from both parents. The phase of 2 variants is most reliably established by testing a parent or child of the proband, but in many cases also can be determined by testing siblings or more distant relatives. In multiplex families, confirming that all affected individuals actually harbor the genotype found in the proband also increases the likelihood that that genotype is truly disease-causing.¹⁸ By reporting such properly segregating genotypes in the literature or through a specially designed database (e.g., vvd.wivr.uiowa.edu), one can strengthen the confidence in those mutations for other physicians caring for other families.

It should go without saying that there is no need to use the same genotyping strategy for every patient. Some phenotypes are so characteristic that they yield a pretest hypothesis that can be evaluated with a single conventional DNA sequencing reaction, which costs less than \$20 in the research setting (see “Methods”). Other phenotypes are associated with a small number of genes that still can be analyzed more quickly and with less financial and statistical cost than an entire exome or genome sequence would incur (Figs 4 and 5).

In this study, we divided all of the inherited eye diseases seen by a single clinician over a 5.5-year period into 62 different categories, and for all but 7 of these categories, we were able to devise very focused tests that cost less than an entire exome to perform. We reserved whole-exome sequencing for the few clinical categories that were too broad for focused screening and for the cases that showed negative results after the initial test. We reserved whole-genome sequencing for 4 families that had a phenotype that strongly implicated a single gene (e.g., patient B in Fig 7), but had no mutations in the coding sequences of that gene. Although this tiered approach resulted in some patients having 2 or even 3 molecular evaluations, the focused tests were so inexpensive—less than half the cost of an exome on average—that the tiered strategy was overall less expensive than it would have been if we performed whole-exome sequencing on every patient (Fig 4). The very customized nature of the prescreening tests also allowed very challenging portions of the genome to be analyzed successfully, such as the highly repetitive portion of exon 15 in *RPGR* that is uninterpretable with most NGS methods. As a result, the sensitivity of our current tiered approach is 6.1% higher than an entirely whole-exome sequencing strategy would be.

Although the tiered strategy currently is 17.7% less expensive overall in our hands than an entire whole-exome sequencing approach would be in the same laboratory using the same personnel, this modest overall cost savings is not the main reason that we use or recommend this approach. The main reasons are to keep the average FGR as low as possible and to detect important disease-causing mutations that otherwise would be missed (patients A–C in Figs 6–8). The clinical pretest decision making necessary to achieve the low FGR results in a very low test cost for a large fraction of the patients (Fig 4). This savings in reagent cost and laboratory bandwidth then can be used to pursue much more expensive investigations, such as cloning the repetitive region of *RPGR* exon 15, in the subset of patients that need it. This results in a higher overall sensitivity of the strategy at a lower cost. It is important to note that as the cost of whole-exome sequencing and the associated analysis continues to fall, it will not supplant the value of specific prescreening tests for many clinical categories until the whole-exome sequencing cost falls to less than that of a single Sanger sequencing reaction. For 13 of the 25 families with clinical Bardet-Biedl syndrome in this cohort, their mutations were found in *BBS1*, and all 13 of these harbored at least 1 M390R allele. As a result, we recommend performing a Sanger sequencing reaction in search of this mutation in all Bardet-Biedl syndrome patients before

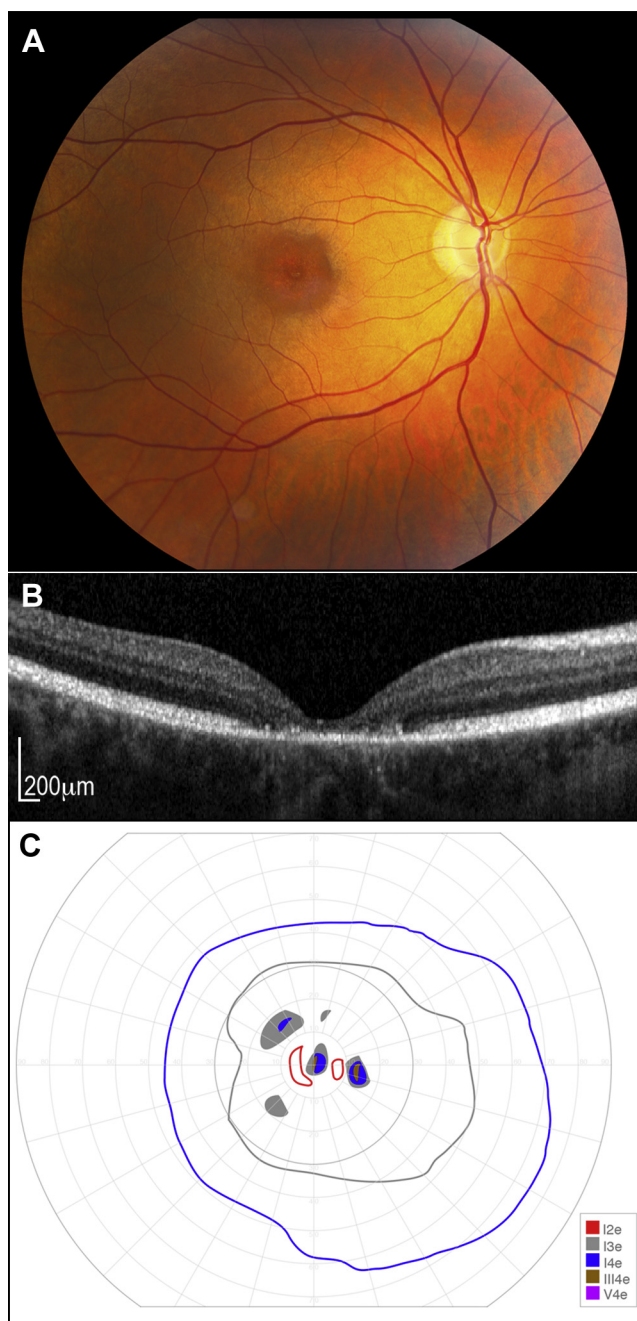


Figure 6. Images from a 47-year-old man with *RPGR*-associated X-linked cone dystrophy. **A**, Fundus photograph of the right eye. **B**, Optical coherence tomography image of the right eye. **C**, Goldmann visual field of the right eye.

proceeding to whole-exome sequencing until the total cost of the latter falls to less than \$50.

It is interesting to consider what would happen to the data shown in Figures 4 and 5 if the research cost of whole-exome and whole-genome sequencing became one tenth what it is today (i.e., \$120 and \$245 per person, respectively). At these price points, the cost of the sequencing would be dwarfed by the cost of the sample handling,

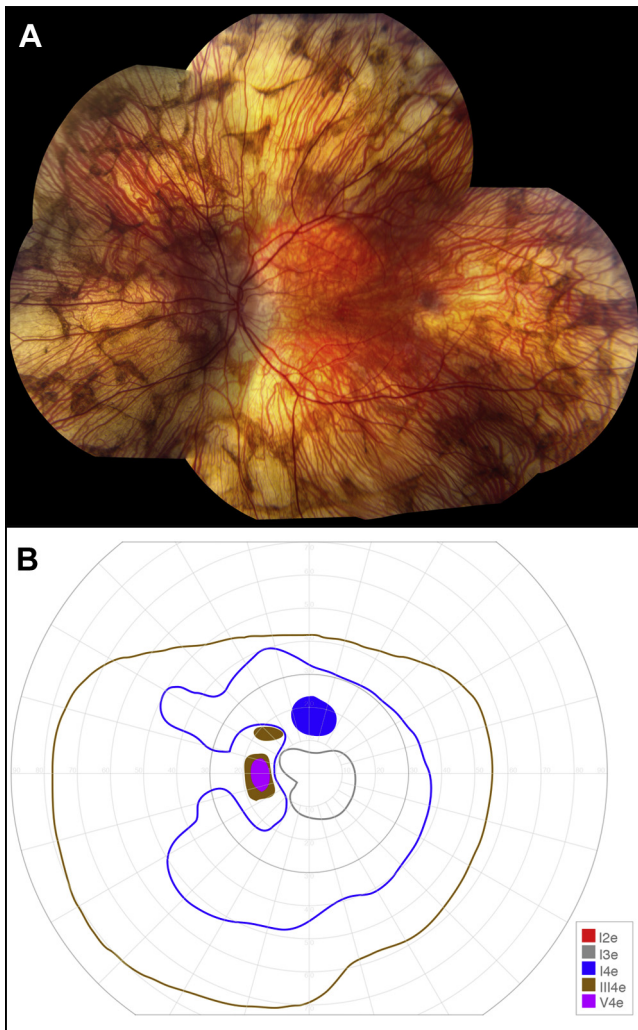


Figure 7. Images from a 8-year-old boy with choroideremia. **A**, Fundus photograph of the left eye. **B**, Goldmann visual field of the left eye.

quality control measures, bioinformatic analysis, report writing, and genetic counseling. As a result, our ratio of whole-exome sequencing to whole-genome sequencing likely would be the inverse of what it is today, and we also would perform many fewer prescreening tests. Most of the latter would be performed to cover the low-complexity parts of the genome that will continue to elude scrutiny by NGS methods. The overall sensitivity of the testing strategy would increase by a few percent because whole-genome sequencing is better at detecting copy number variations. However, the need for a narrow pretest hypothesis would be identical to the need today because the average FGR per base pair of investigated genome is an immutable fact of nature that is completely unaffected by the costs of the methods we use or the speed with which we use them.

One might expect that our next step in studying the cohort presented in this article would be to perform whole-genome sequencing in the 240 families that have yet to have their disease-causing mutations identified. However, it is important to note that these families harbor an average of

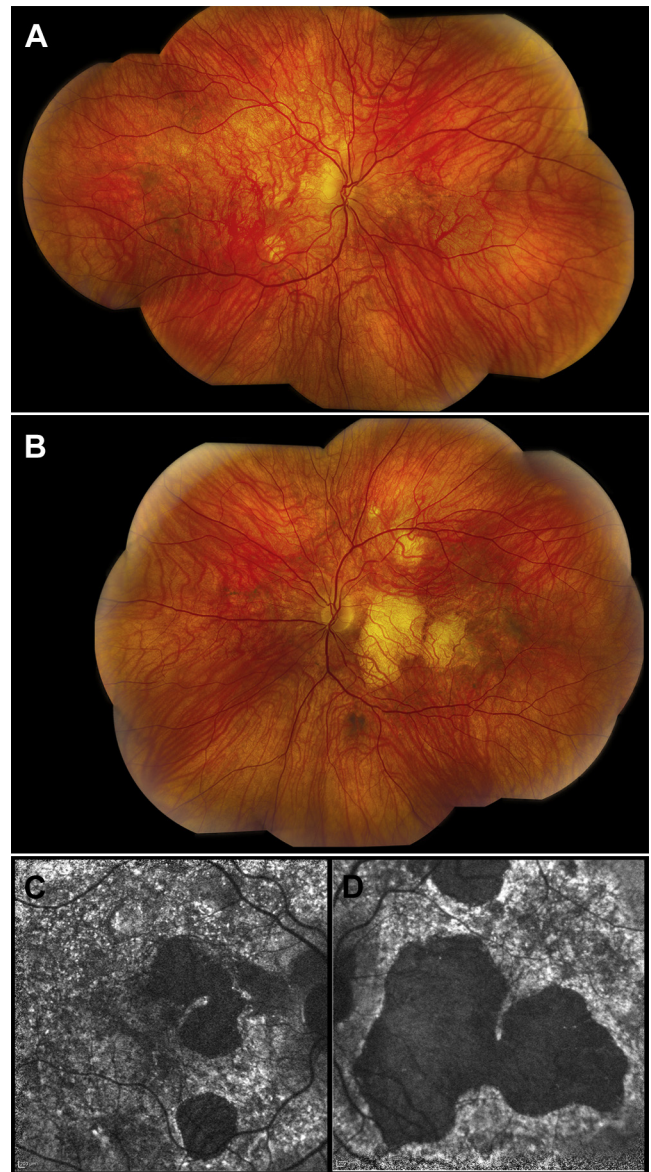


Figure 8. Images from a 48-year-old woman with maternally inherited diabetes and deafness. **A**, Fundus photograph of the right eye. **B**, Fundus photograph of the left eye. **C**, **D**, Fundus autofluorescence images of (C) the right eye and (D) the left eye.

16.5 plausible disease-causing mutations among the 305 candidate genes we considered in this study (Fig 1). It seems most likely to us that most of the genotypes remaining to be discovered in this cohort lie at least in part among the coding sequence variations that we already have detected or the coding sequences of other genes and that further clinical investigation of these families is likely to be more fruitful than increasing the number of rare variants to consider by more than 2 orders of magnitude. The aggressive ascertainment of additional members of these 240 families will allow us to strengthen or rule out many of the plausible disease-causing variants we already have identified on the basis of their segregation within the families. Continued scrutiny of the families with positive results

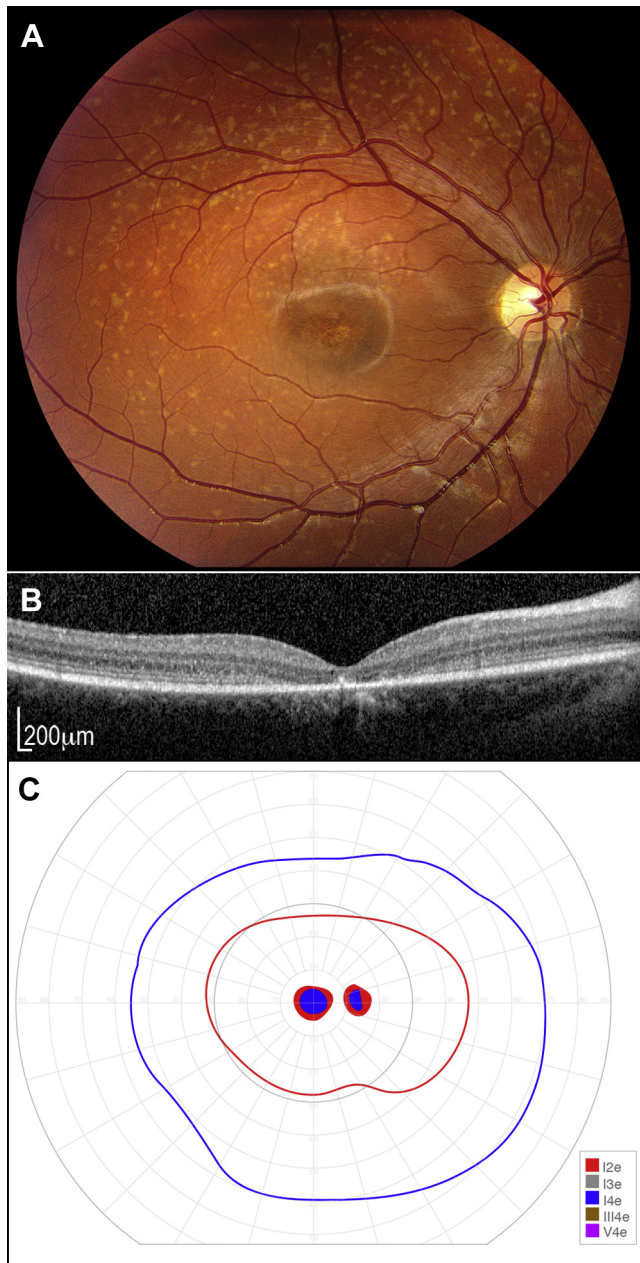


Figure 9. Images from a 10-year-old girl with ABCA4-associated Stargardt disease. **A**, Fundus photograph of the right eye. **B**, Optical coherence tomography image of the right eye. **C**, Goldmann visual field of the right eye.

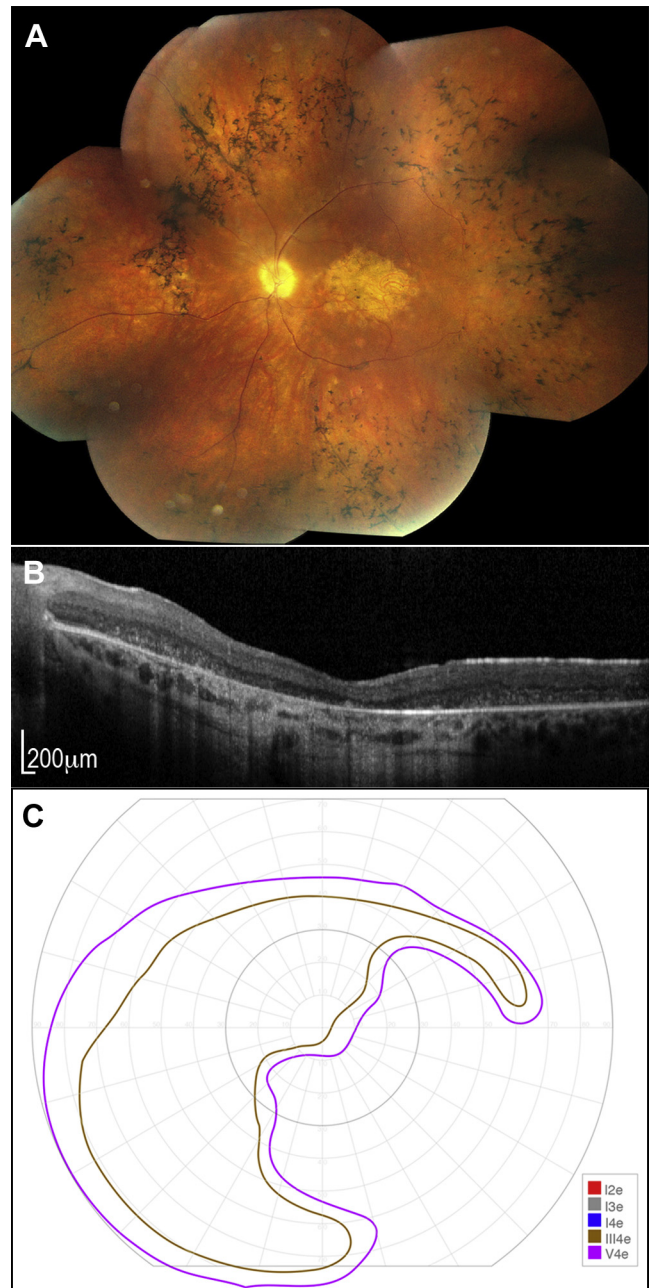


Figure 10. Images from a 42-year-old man with ABCA4-associated Stargardt disease. **A**, Fundus photograph of the left eye. **B**, Optical coherence tomography image of the left eye. **C**, Goldmann visual field of the left eye.

in this cohort also may reveal some characteristic clinical features that would favor a specific 1-allele hypothesis sufficiently such that whole-genome sequencing would be indicated in that family. This *families first* strategy would not change even if the cost of whole-genome sequencing fell 10-fold. As noted above, the reason for this is that the amount of normal genetic variation in the genome is extremely large and independent of sequencing cost. The most powerful resources for overcoming this noise are, and will continue to be, large and well-characterized patient resources.^{28,33}

The disadvantages of a tiered testing strategy are that it requires very accurate communication between the clinic and the laboratory to gain the benefits described in this study and that tiered tests take much longer to perform than fragment capture panels. Although there are few situations in which a 3- or 4-month difference in testing time is clinically significant for a patient with a slowly progressive retinal degeneration, it is unquestionable that many families are anxious to have the cause of their disease identified as quickly as possible.

The keys to keeping the FGR down and sensitivity high are (1) to make the best clinical diagnosis possible before ordering a genetic test and to use this diagnosis to choose the simplest test that is likely to yield a finding for that diagnosis, (2) to obtain samples from parents and siblings of simplex families and as many affected individuals as possible from multiplex pedigrees for use in evaluating the results in the proband, (3) to know the cost breakpoint between multiple focused prescreens and whole-exome sequencing and to switch to whole-exome sequencing before exceeding that breakpoint, and (4) to take advantage of the slow progression of most of these diseases by trying to have a result for the patient at their next visit, instead of some arbitrarily short turnaround time that will inflate the cost of the test artificially.

Although the 76% sensitivity achieved in this study is a far cry from the 0% of 1986, it is likely to become even higher as we continue to analyze the 240 probands of this cohort whose molecular pathophysiologic features have yet to be determined. Some of the probands in this cohort are likely to have had inflammatory insults to their retinas that mimic Mendelian disease, and it is possible that a predisposition to such disease may be detectable in their DNA as our knowledge of the genetics of the immune system continues to expand. There will certainly be additional disease-causing genes identified in the future by subjecting cohorts like this one to more sophisticated analysis or by studying multiplex families who lack mutations in currently known genes. There are also likely to be additional examples of nonexomic^{29,33,46} and mitochondrial disease discovered, as well as convincing cases of multiple genes interacting with one another to cause disease.^{47,48}

One advantage that we have today over 1986 is the ability to perform many genetic tests recursively, in silico, using inexhaustible data that are stored on servers instead of exhaustible DNA stored in freezers. Another advantage is the ability to derive phenotypically accurate retinal cell cultures from accessible tissues like skin and to use these cells to test hypotheses that are generated from the DNA analysis.^{29,40,46} However, the most valuable resources needed to make these new discoveries are unchanged from 1986: relatively large numbers of patients with exceptionally detailed clinical information and large numbers of affected and unaffected family members that can be used to evaluate the many hypotheses that arise when studying the probands. As a result, the astute clinician, who is a good observer and record keeper and who is willing to do whatever is necessary to find the correct answer and an effective treatment for his or her patient, remains the most valuable component of the entire effort.

References

1. Cavenee WK, Dryja TP, Phillips RA, et al. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature*. 1983;305:779-784.
2. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*. 1986;323:643-646.
3. Dryja TP, McGee TL, Reichel E, et al. A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature*. 1990;343:364-366.
4. Maguire AM, Simonelli F, Pierce EA, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med*. 2008;358:2240-2248.
5. MacLaren RE, Groppe M, Barnard AR, et al. Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial. *Lancet*. 2014;383:1129-1137.
6. Constable IJ, Blumenkranz MS, Schwartz SD, et al. Gene therapy for age-related macular degeneration. *Asia Pac J Ophthalmol (Phila)*. 2016;5:300-303.
7. Ran FA, Hsu PD, Lin C-Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154:1380-1389.
8. Jinek M, East A, Cheng A, et al. RNA-programmed genome editing in human cells. *Elife*. 2013;2:e00471.
9. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823-826.
10. Wiley LA, Burnight ER, Songstad AE, et al. Patient-specific induced pluripotent stem cells (iPSCs) for the study and treatment of retinal degenerative diseases. *Prog Retin Eye Res*. 2015;44:15-35.
11. Wiley LA, Burnight ER, DeLuca AP, et al. cGMP production of patient-specific iPSCs and photoreceptor precursor cells to treat retinal degenerative blindness. *Sci Rep*. 2016;6:30742.
12. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860-921.
13. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291:1304-1351.
14. Cideciyan AV, Swider M, Aleman TS, et al. Reduced-illumination autofluorescence imaging in ABCA4-associated retinal degenerations. *J Opt Soc Am A Opt Image Sci Vis*. 2007;24:1457-1467.
15. McCulloch DL, Marmor MF, Brigell MG, et al. ISCEV standard for full-field clinical electroretinography (2015 update). *Doc Ophthalmol*. 2015;130:1-12.
16. Stone EM. Leber congenital amaurosis—a model for efficient genetic testing of heterogeneous disorders: LXIV Edward Jackson Memorial Lecture. *Am J Ophthalmol*. 2007;144:791-811.
17. Newton CR, Graham A, Heptinstall LE, et al. Analysis of any point mutation in DNA. The Amplification Refractory Mutation System (ARMS). *Nucleic Acids Res*. 1989;17:2503-2516.
18. Stone EM. Finding and interpreting genetic variations that are important to ophthalmologists. *Trans Am Ophthalmol Soc*. 2003;101:437-484.
19. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589-595.
20. 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.
21. Krumm N, Sudmant PH, Ko A, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res*. 2012;22:1525-1532.
22. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2016;32:1220-1222.
23. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285-291.
24. Lott MT, Leipzig JN, Derbeneva O, et al. mtDNA variation and analysis using Mitomap and Mitomaster. *Curr Protoc Bioinformatics*. 2013;44:1.23.1-1.23.26.
25. Weleber RG, Michaelides M, Trzupke KM, et al. The Phenotype of Severe Early Childhood Onset Retinal

- Dystrophy (SECORD) from mutation of RPE65 and differentiation from Leber congenital amaurosis. *Invest Ophthalmol Vis Sci.* 2011;52:292-302.
26. Genead MA, Fishman GA, Stone EM, Allikmets R. The natural history of Stargardt disease with specific sequence mutation in the ABCA4 gene. *Invest Ophthalmol Vis Sci.* 2009;50:5867-5871.
 27. United States Census Bureau. Population by sex and selected age groups: 2000 and 2010. <https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf>; 2010 Accessed January 7, 2017.
 28. Braun TA, Mullins RF, Wagner AH, et al. Non-exonic and synonymous variants in ABCA4 are an important cause of Stargardt disease. *Hum Mol Genet.* 2013;22:5136-5145.
 29. Tucker BA, Mullins RF, Streb LM, et al. Patient-specific iPSC-derived photoreceptor precursor cells as a means to investigate retinitis pigmentosa. *Elife.* 2013;2:e00824.
 30. Khan SY, Ali S, Naeem MA, et al. Splice-site mutations identified in PDE6A responsible for retinitis pigmentosa in consanguineous Pakistani families. *Mol Vis.* 2015;21:871-882.
 31. Bellingham J, Davidson AE, Aboshiha J, et al. Investigation of aberrant splicing induced by AIPL1 variations as a cause of Leber congenital amaurosis. *Invest Ophthalmol Vis Sci.* 2015;56:7784-7793.
 32. den Hollander AI, Koeneke RK, Yzer S, et al. Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. *Am J Hum Genet.* 2006;79:556-561.
 33. Small KW, DeLuca AP, Whitmore SS, et al. North Carolina macular dystrophy is caused by dysregulation of the retinal transcription factor PRDM13. *Ophthalmology.* 2016;123:9-18.
 34. Howell N, Bindoff LA, McCullough DA, et al. Leber hereditary optic neuropathy: identification of the same mitochondrial ND1 mutation in six pedigrees. *Am J Hum Genet.* 1991;49:939-950.
 35. van den Ouweland JM, Lemkes HH, Gerbitz KD, Maassen JA. Maternally inherited diabetes and deafness (MIDD): a distinct subtype of diabetes associated with a mitochondrial tRNA(Leu)(UUR) gene point mutation. *Muscle Nerve Suppl.* 1995;3:S124-S130.
 36. Vervoort R, Lennon A, Bird AC, et al. Mutational hot spot within a new RPGR exon in X-linked retinitis pigmentosa. *Nat Genet.* 2000;25:462-466.
 37. Ayyagari R, Demirci FY, Liu J, et al. X-linked recessive atrophic macular degeneration from RPGR mutation. *Genomics.* 2002;80:166-171.
 38. Yang Z, Peachey NS, Moshfeghi DM, et al. Mutations in the RPGR gene cause X-linked cone dystrophy. *Hum Mol Genet.* 2002;11:605-611.
 39. Schindler EI, Nylen EL, Ko AC, et al. Deducing the pathogenic contribution of recessive ABCA4 alleles in an outbred population. *Hum Mol Genet.* 2010;19:3693-3701.
 40. Tucker BA, Scheetz TE, Mullins RF, et al. Exome sequencing and analysis of induced pluripotent stem cells identify the cilia-related gene male germ cell-associated kinase (MAK) as a cause of retinitis pigmentosa. *Proc Natl Acad Sci U S A.* 2011;108:E569-E576.
 41. Fromer M, Purcell SM. Using XHMM software to detect copy number variation in whole-exome sequencing data. *Curr Protoc Hum Genet.* 2014;81:7.23.1-7.23.21.
 42. Suzuki T, Tsurusaki Y, Nakashima M, et al. Precise detection of chromosomal translocation or inversion breakpoints by whole-genome sequencing. *J Hum Genet.* 2014;59:649-654.
 43. Wildschutte JH, Baron A, Diroff NM, Kidd JM. Discovery and characterization of Alu repeat sequences via precise local read assembly. *Nucleic Acids Res.* 2015;43:10292-10307.
 44. Merico D, Roifman M, Braunschweig U, et al. Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman syndrome by disrupting minor intron splicing. *Nat Commun.* 2015;6:8718.
 45. Käsäkoski J, Jääskeläinen J, Jääskeläinen T, et al. Complete androgen insensitivity syndrome caused by a deep intronic pseudoexon-activating mutation in the androgen receptor gene. *Sci Rep.* 2016;6:32819.
 46. Tucker BA, Cranston CM, Anfinson KA, et al. Using patient-specific induced pluripotent stem cells to interrogate the pathogenicity of a novel retinal pigment epithelium-specific 65 kDa cryptic splice site mutation and confirm eligibility for enrollment into a clinical gene augmentation trial. *Transl Res.* 2015;166:740-749.e1.
 47. Dryja TP, Hahn LB, Kajiwarra K, Berson EL. Dominant and digenic mutations in the peripherin/RDS and ROM1 genes in retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 1997;38:1972-1982.
 48. Kajiwarra K, Berson EL, Dryja TP. Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. *Science.* 1994;264:1604-1608.

Footnotes and Financial Disclosures

Originally received: February 15, 2017.

Final revision: April 5, 2017.

Accepted: April 6, 2017.

Available online: May 27, 2017.

Manuscript no. 2017-339.

¹ Department of Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine, Iowa City, Iowa.

² Stephen A. Wynn Institute for Vision Research, the University of Iowa, Iowa City, Iowa.

³ Department of Pediatrics, University of Iowa Carver College of Medicine, Iowa City, Iowa.

Financial Disclosure(s):

The author(s) have no proprietary or commercial interest in any materials discussed in this article.

Supported by the National Eye Institute, Bethesda, Maryland (grant nos.: R01EY024588 and R01EY026008); the Wynn Institute for Vision Research, Iowa City, Iowa; the Carver Non-Profit Genetic Testing Laboratory; the Stephen A. Wynn Foundation, Iowa City, Iowa; and Mark J. Daily, MD.

Author Contributions:

Conception and design: Stone

Analysis and interpretation: Stone, Andorf, Whitmore, DeLuca, Giacalone, Streb, Braun, Mullins, Scheetz, Sheffield, Tucker

Data collection: Stone, Andorf, Whitmore, DeLuca, Giacalone, Streb, Braun, Mullins, Scheetz, Sheffield, Tucker

Obtained funding: Stone, Tucker, Mullins

Overall responsibility: Stone, Andorf, Whitmore, DeLuca, Mullins, Scheetz, Sheffield, Tucker

Abbreviations and Acronyms:

AD = autosomal dominant; **ADNIV** = autosomal dominant neovascular inflammatory vitreoretinopathy; **AR** = autosomal recessive; **AR-1** = autosomal recessive-1 allele identified; **ARMS** = amplification refractory mutation system; **AZOOR** = acute zonal occult outer retinopathy; **BBS** = Bardet-Biedl syndrome; **BWA** = Burrows Wheeler aligner; **CRISPR** = clustered regularly interspaced short palindromic repeats; **CSNB** = congenital stationary night blindness; **CSSD** = congenital stationary synaptic dysfunction; **DDND** = developmental delay and/or neuromuscular degeneration; **ECORD** = early childhood onset retinal dystrophy; **EPP** = estimate of

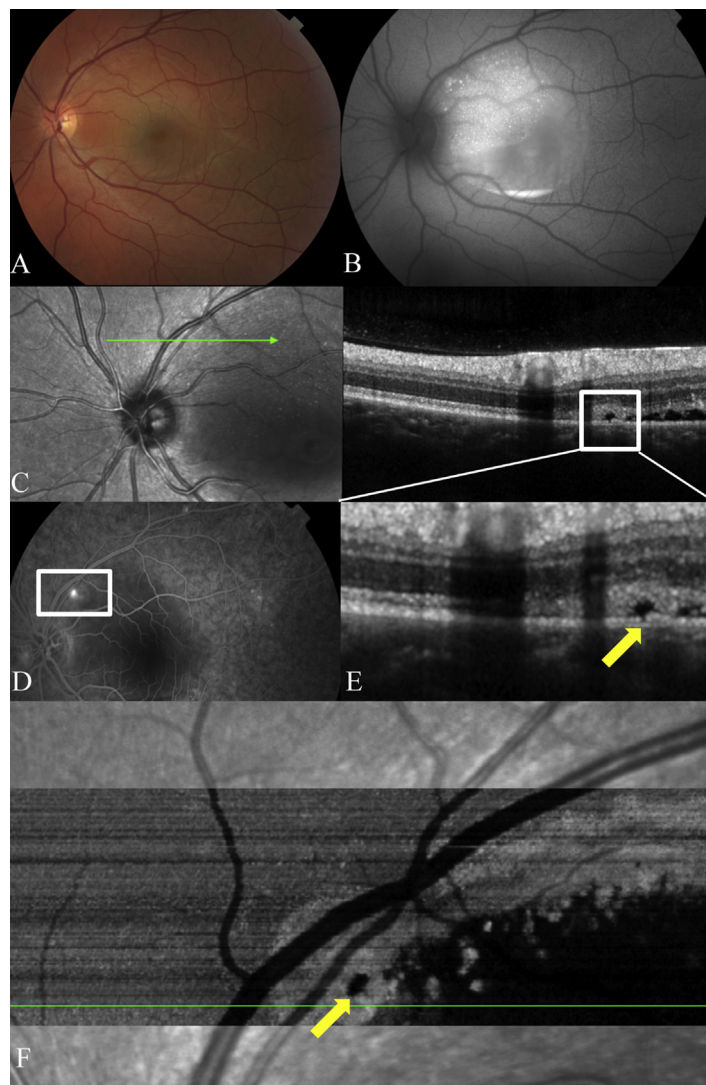
pathogenic probability; **ERG** = electroretinogram; **EV** = erosive vitreoretinopathy; **ExAC** = Exome Aggregation Consortium; **FEVR** = familial exudative vitreoretinopathy; **FGR** = false genotype rate; **GATK** = genome analysis toolkit; **HMA** = homocystinuria with macular atrophy; **HPCD** = helicoid peripapillary chorioretinal degeneration; **ISCEV** = International Society for Clinical Electrophysiology of Vision; **IVS** = intervening sequence; **LB** = lysogeny broth; **L/M Opsin** = long/medium wave length opsin; **LCA** = Leber congenital amaurosis; **LCHAD** = long-chain 3-hydroxyacyl-coenzyme A dehydrogenase deficiency; **LHON** = Leber hereditary optic neuropathy; **MCLMR** = microcephaly congenital lymphedema and chorioretinopathy; **MDPD** = mutation detection probability distribution;

MIDD = maternally inherited diabetes and deafness; **MIS** = missense; **Mito** = mitochondrial; **NGS** = next-generation sequencing; **PCR** = polymerase chain reaction; **PV** = plausible variants; **RP** = retinitis pigmentosa; **RPE** = retinal pigment epithelium; **SECORD** = severe early childhood onset retinal dystrophy; **TA Cloning** = thymine and adenine cloning; **TERM** = terminating; **VVD** = vision variation database; **XL** = X-linked.

Correspondence:

Edwin M. Stone, MD, PhD, The Stephen A. Wynn Institute for Vision Research, 375 Newton Road, 4111 MERF, Iowa City, IA 52242. E-mail: edwin-stone@uiowa.edu.

Pictures & Perspectives



Multimodal Imaging in Central Serous Chorioretinopathy

A 29-year-old man presented with reduced vision in his left eye to 20/60 for 3 months. Figure 1A shows a serous detachment with subretinal dots and dependent vitelliform material. Fundus autofluorescence (Fig 1B) shows hyperautofluorescence of both the dots and the vitelliform material. Fluorescein angiography (Fig 1D) shows a focal leak corresponding to a hyporeflective subretinal lucency seen on optical coherence tomography (OCT; Fig 1C). A micro-rip in the retinal pigment epithelium can be seen with magnification (Fig 1E). En face OCT (Fig 1F) shows a hyporeflective area above the focal leak (yellow arrows) (Magnified version of Fig 1A-F available at www.aaojournal.org).

SARRA GATTOUSSI, MD
K. BAILEY FREUND, MD

Vitreous Retina Macula Consultants of New York, New York, New York