The current status of molecular diagnosis of inherited retinal dystrophies

John (Pei-wen) Chiang and Karmen Trzupek

Purpose of review
We are witnessing lightning-fast advances in the molecular diagnosis of inherited retinal dystrophies, mainly due to the widespread use of next-generation sequencing technologies. The purpose of this review is to highlight the breadth of findings from this in-depth testing approach, and to propose changes to our traditional testing and diagnostic paradigms. Lessons learned from modern molecular testing suggest that the previous concept of inherited retinal dystrophies as a group of ‘single gene diseases’ may require a significant update.

Recent findings
All of the known retinal dystrophies genes can now be sequenced. In many cases, this nonhypothesis driven testing strategy is uncovering mutations in unsuspected genes, generating data that challenges established concepts of genetic mechanisms and provides insights regarding genes previously thought to be exclusively related to syndromic disease. Recent advances in testing have improved not only the breadth, but also the depth of genetic data. For example, deep intronic sequencing has uncovered many novel intronic mutations/variations in the ABCA4 gene.

Summary
Currently, in approximately 50–60% of patients with nonsyndromic retinal dystrophy, the disease mechanism can be identified. The presence of pathogenic alleles in more than one gene is not uncommon. Retinal dystrophy, with relatively defined clinical presentations and a large but limited number of genes involved, is becoming a model for the next-generation study of molecular disease mechanisms.

Keywords
Genetic testing, molecular diagnosis, mutation, retinal dystrophy, retinitis pigmentosa

INTRODUCTION
Retinitis pigmentosa is the most common form of inherited retinal dystrophy, affecting one in 3500 to one in 4000 people in the USA and Europe (Genetic Home Reference, http://ghr.nlm.nih.gov/condition/retinitis-pigmentosa). Increasingly, molecular diagnosis is becoming an integral part of clinical diagnosis. Variable disease expression, genotypic heterogeneity (one phenotype caused by more than one gene), phenotypic heterogeneity (different mutations in one gene resulting in different phenotypes), reduced penetrance, and the progressive nature of these diseases complicate clinical diagnosis. Molecular testing has long aided diagnosis, but unclear inheritance patterns in families and the large number of genes involved has traditionally complicated the task of choosing the appropriate genes to test. With the arrival of next-generation sequencing (NGS), however, large numbers of genes can be sequenced simultaneously. At the same time, potential modifiers can be unraveled. Molecular testing is gradually becoming the gold standard in clarifying diagnoses and elucidating the underlying genetic mechanisms.

Before whole genome sequencing becomes the universal testing for every genetic condition, the current testing strategies are based on either single gene testing or panel testing. Cost and confidence of making accurate clinical diagnosis for conditions with strong phenotype and genotype correlations are important factors to be considered when...
Ocular genetics

KEY POINTS

- Next-generation sequencing offers unprecedented opportunities especially to conditions such as retinal dystrophy.
- Retinal dystrophy, with its main clinical presentation specific to retina and its limited number of, but well characterized, genes involved is becoming a model for the study of rare genetic conditions.
- For the first time, almost the entire genes involved can be sequenced simultaneously and specifically with deep coverage. Comprehensive and large scale sequencing efforts are changing our understanding of genetic mechanisms.
- Molecular diagnosis of retinal dystrophy is becoming an integral part of clinical diagnosis and it may even become the first line of diagnostic tool in the near future.

ordering single gene testing. X-linked juvenile retinoschisis (RS1 gene), Stargardt disease (ABCA4 gene), Choroideremia (CHM gene), and Vitelliform macular dystrophy (PRPH2/RDS or BEST1 genes) are the conditions with more single testing. In contrast, a condition such as retinitis pigmentosa is more ideal to be tested by a large gene panel. One additional advantage for panel testing is that modifiers, digenic mutations, and multiallelic interactions can also be identified through panel testing. Taken together, testing by a comprehensive panel is more ideal for retinal dystrophy.

ABCA4, the gene which when mutated can result in Stargardt disease, is one of the most studied ocular disease genes and also one of the most commonly mutated genes in retinal dystrophies. We will use this as an example of how modern sequencing strategies can be applied.

Stargardt disease affects, approximately, one in 8000 to one in 10000 individuals (Genetic Home Reference, http://ghr.nlm.nih.gov/condition/stargardt-macular-degeneration). With its distinct clinical presentation, relative high disease incidence and high mutation concentration in a single gene, ABCA4 can become a paradigm for the study of retinal dystrophy disease mechanisms. However, mutations in ABCA4 can present with a wide array of phenotypes. Several interesting aspects of ABCA4 molecular genetics are worth discussing:

Correlation of phenotypes with mutation severities

The correlation of phenotypes with mutation severity in ABCA4 has long been proposed [1,2]. Mutation detection rate also appears to correlate with age of onset. Younger patients usually have two or more mutations identified in ABCA4 [3,4]. A quantitative threshold effect of mutation load is likely at work, which may conflict with the traditional concept of single gene Mendelian inheritance. Mutations in ABCA4 can cause Stargardt, cone or cone–rod dystrophy and retinitis pigmentosa, with Stargardt disease frequently harboring less severe mutations [5]. In many cases, however, no clear genotype–phenotype correlations emerge. The presence of genetic modifiers may be the missing link. Recently, a rare combination of mutations in ABCA4 and GRM6 was reported in a patient with atypical Stargardt disease and an unusual electroretinography (ERG) phenotype more consistent with congenital stationary night blindness (CSNB) [6]. Two mutations in ABCA4 and two mutations in GRM6 were identified in the patient. Although this specific combination may be extremely rare, the identification of pathogenic variants in additional retinal disease genes is common when a large number of genes are sequenced. Historically, clinicians and researchers stop testing when the presumed disease-causing mutation(s) is identified. With NGS technologies, secondary mutations are increasingly being identified.

Deep intronic variants in ABCA4

Numerous reports have confirmed that approximately 15% of patients with a clinical diagnosis of Stargardt disease have only one identifiable mutation in the coding region of ABCA4. Two recent studies shed light on some of the missing mutations [7**]. In the first report, five deep intronic mutations in ABCA4 were identified, increasing the detection rate of finding two mutations to 91.8% in the studied cohort [7**]. In the second report [8**], 114 patients with only one previously known ABCA4 mutation were studied. Deep intronic mutations were identified in an additional 27/114 (23.7%); no second mutations were identified in 36/114 (31.6%) and intronic variants of unknown significance were identified in the remaining 51/114 (44.7%) patients. The increase in mutation detection by incorporating these five deep intronic mutations was confirmed in two other studies [9,10]. However, the prediction and confirmation of pathogenicity for the additional intronic variants has proven difficult, especially when the variant is rare [8**]. The detection of rare deep intronic variants in a highly polymorphic gene such as ABCA4 is not uncommon. Some of these deep intronic variants likely do not directly contribute to the patient’s disease.
Looking in the wrong place?

Even following whole gene and deep intronic sequencing, some patients have only one identifiable ABCA4 mutation. Do some of these patients actually have a mutation in a different gene? The carrier rate of ABCA4 mutations is approximately 2–3% in the general population. Also, we cannot rule out the combined effect of two mutations in two different retinal disease genes, normally functioning synergistically. One of our recent cases illustrates this complexity. A patient with the clinical diagnosis of simplex cone dystrophy diagnosed approximately age 10 was tested using our Retinal Dystrophy SmartPanel v2 (154 genes). Two variants were identified in the ABCA4 gene: p.T829M: c.2486C>T (rs139250920; predicted to be benign by PolyPhen-2) and c.4253+43G>A (rs61754045). The p.T829M was reported as a mutation in the literature once. The clinical significance of the c.4253+43G>A variant is unclear, and could be a hypomorphic allele. However, patients with cone or cone-rod dystrophy typically have more severe mutations in ABCA4 [5]. In addition to the ABCA4 variants, a heterozygous frameshift mutation (c.6120delC) was also identified in GPR179—a gene associated with recessive complete CSNB. As the disease mechanism remained unclear, whole exome sequencing was later performed and a heterozygous deletion of the entire CRX gene was identified. Unfortunately, parental samples were not available to determine whether the CRX deletion was de novo in the patient. Large deletions of CRX are known to be associated with dominant de-novo cases of Leber Congenital Amaurosis and early-onset cone dystrophy. This case may raise more questions than it answers: are the ABCA4 variants a ‘red herring’, unrelated to the patient’s disease? Could ABCA4 be a modifier of disease expression in this case? At what point is clinical sequencing complete in retinal dystrophy patients with positive results? In patients with only one identifiable mutation in ABCA4, after sequencing the entire coding regions and the reported deep intronic mutations, how many actually possess additional mutation(s) in different genes? A systematic and comprehensive mutation detection strategy may be warranted for this group of patients.

A powerful tool in assisting variant calling

With the progress of several large scale whole exome and whole genome-sequencing projects, allele frequency for a significant portion of variants are now well characterized (http://exac.broadinstitute.org/). Variations from 60,706 unrelated individuals have been recorded and allele frequency from different races are shown for each variant. Interestingly, this new tool has also generated uncertainties to some published mutations. For example, allele frequency for the common Whites ABCA4 mutation p.G1961E (rs1800553) is 0.004723 in Europeans (non-Finnish) but it is 0.01498 in South Asians. Typically, allele frequency cutoff is set at 0.005 for a rare genetic disease. By this standard, a variant with an allele frequency of 0.01498 is not called as a mutation, assuming disease incidence is not drastically higher in the population. Similarly, p.R2107H (rs62642564) in ABCA4 is also a reported mutation. Allele frequency in Europeans (non-Finnish) is 7.506e-05 but it is 0.02002 in Africans. The biggest challenge is apparently to some reported dominant mutations. For example, the GUCA1A p.P50L (rs104893968) variant was reported as a dominant mutation. However, allele frequency for this variant is 0.00103. Based on our family studies, this variant has been identified in normal family members. This variant, if truly pathogenic, likely requires additional mutations from other genes to cause disease. Clearly, evaluation of dominant mutations should be more cautious. Allele frequency of dominant mutations can be instructive in the evaluation of pathogenicity [11]. Allele frequency of any dominant mutation should be less than the disease incidence. Pathogenicity of some of the published dominant mutations are in question. Multiallelic interactions may be considered in some cases.

Next-generation sequencing

As little as a few years ago, RHO was probably the most commonly sequenced gene for patients with retinitis pigmentosa. Mutation detection rate was poor for patients with simplex or presumed autosomal recessive retinitis pigmentosa, and there was little clinical utility of genetic testing. Presumed inheritance type guided molecular testing, and typically only a few genes were tested for any given patient. This approach is especially problematic for conditions such as retinal dystrophies, wherein at least 200 disease-associated genes are known (https://sph.uth.edu/retnet/) and genotypic and phenotypic heterogeneity makes it difficult to predict genotype (and thus the gene to select for testing) based on phenotype. Our past narrow interpretation of inheritance modes and mutation mechanisms will require some, if not significant, modifications. When large number of genes are sequenced, some patients have identified possibly pathogenic variants in multiple genes or mutations in a single gene that were not expected based on phenotype. The presence of multiple presumed
pathogenic variants in multiple genes with various inheritance modes can complicate interpretation of disease mechanisms. The hallmarks of dominant inheritance – variable expression, variable age of onset, and penetrance may be explained in the future by the presence or absence of additional pathogenic variants.

**The challenges of modern genetic testing**

Despite its advantages, the introduction of clinical NGS testing has created challenges. Because the technologies are new, clinical diagnostic laboratories are still learning and adapting. The advent of NGS testing has identified a large number of variants of uncertain significance. Data interpretation has become a challenge. There is often no consensus on interpretation of sequence variants that are not clearly pathogenic. The initial hope that whole exome sequencing (WES) would replace disease-specific gene panels is gradually fading away. Panel testing is frequently more sensitive than WES testing [12*]. The reason is mainly due to the method utilized in the target enrichment step. Currently, many laboratories use hybridization-based method (capture method) to enrich targets for sequencing. Gene hybridization relies on sequence homology. However, hybridization also enriches DNA from additional regions with homologies such as pseudogenes and gene family members. The ratio of out-of-target captures versus on-target captures vary to different extents depending upon the genes involved. Some genes have pseudogenes and some others do not. Certain genes belong to gene families (with several members of similar genes in the families). Additionally, no specific capture baits can be designed for some repetitive regions. Also hybridization condition is not universal for every target region. Therefore, the target regions can never be evenly represented in the captures. Sequencing depth then becomes critical in coverage. For specific gene panels, the issues raised above are less critical due to less genes are covered.

Equally important, data analysis becomes very demanding when large number of genes are sequenced. How to filter out false-positive variants without filtering out true mutations is essential in order to reduce false-negative rate. The practice of setting sequencing depth at a lower range (such as 5X or 10X depths) in order to increase coverage can be problematic. When sequencing depth is lower, more false-positive variants appears (variants identified when only sequenced five times are less reliable than variants identified by sequencing 100 times).

**Mutation detection rate**

Retinal dystrophies are one of the most heterogeneous group of genetic diseases. In order to increase mutation detection rate, a large number of genes must be sequenced. NGS testing of several retinal dystrophy patient cohorts have been reported recently, with published mutation detection rates for nonsyndromic retinal dystrophy of 37 [13*], 51 [12*], 55 [14], 60 [15], and 70% [16], using targeted enrichment NGS panel tests.

This wide range in mutation detection rate is not necessarily directly attributable to the number of genes analyzed. Eisenberger et al. reported the highest mutation detection rate but the lowest number of included genes. Instead, mutation detection depends upon the makeup of the patient cohort and the inclusion of additional methodologies to increase detection of variants not easily identified with sequencing. The lowest detection rate reported (37%; Wang et al.) was obtained from a cohort of ‘unsettled’ probands with simplex or presumed recessive retinitis pigmentosa, not previously solved using gene-directed sequencing. Laboratory analyses used in these different studies variably included techniques such as copy number variant analysis, long range PCR amplification of ‘hard-to-sequence’ targets to increase coverage of highly repetitive sequences, and mutation-specific analysis of deep intronic mutations. All of these techniques should increase mutation detection in a retinal dystrophy cohort, due to the presence of deep intronic mutations (e.g., ABCA4 and USH2A), highly repetitive regions (e.g., RPGR ORF15), and large deletions (e.g., EYS and USH2A). Still, certain regions in these genes can be missed due to uneven coverage from the capture panel design in these studies.

**The unexpected**

The cost of WES has reduced dramatically over the last several years. Even with broad WES testing, a significant proportion of retinal dystrophy patients still has no mutations identified. The reasons are likely multiple including: novel, as yet undiscovered genes; the potential presence of mutations in genes not previously known to be associated with retinal diseases; the presence of mutations outside the target regions; uneven coverage of target genes; epigenetic mechanisms; inaccurate variant calling, especially of mild alleles in syndromic genes; and uncertain clinical diagnoses. The identification of CLN3, which when mutated was previously known to cause a form of neuronal ceroid lipofuscinosis, as a nonsyndromic retinal dystrophy gene is the latest example of a syndromic gene getting reclassified as a
Molecular diagnosis of inherited retinal dystrophies

Chiang and Trzupka

Retinal dystrophies have long proven to be a challenge for molecular diagnosis. Overlapping clinical presentations leading to uncertainty in clinical diagnosis, unclear inheritance patterns, allele heterogeneity, variable ages of onset, and the huge number of genes involved all made molecular diagnosis extremely difficult in the past. However, the arrival and the broad application of NGS has revolutionized molecular diagnosis of retinal dystrophy. In fact, retinal dystrophies now have a much higher molecular detection rate than some other conditions, such as autism and most forms of inherited cardiomyopathy. Also, it is very likely that most of the common retinal dystrophy genes have been identified. For the first time, all of the known retinal dystrophy genes can now be sequenced for less than US$3000. This sequencing approach currently makes more sense because it offers higher coverage, more confidence in data interpretation, lower cost, and better clinical utility. New phenotype and genotype correlations are beginning to emerge from the vast amount of data generated. Traditional concepts of inheritance are being challenged. Further, some reported mutations in the past are now being revisited in light of a richer allele frequency database. Syndromic genes are being considered as candidate genes for apparently nonsyndromic disease and vice versa. Quantitative threshold effect and milder mutation alleles may be in play. Finally, it will take collaboration throughout the genetics community to coordinate variant calling and data sharing. Despite all these achievements, many patients can still not access commercial genetic testing, particularly given that insurance companies do not always cover genetic testing for rare conditions. Making genetic testing in the less developed nations, where the burden of genetic disease may be even higher, remains a significant challenge. One of the greatest hurdles moving forward will be not just to improve testing methodologies, but to lower testing costs. Our greatest challenge may be democratizing genetic testing to ensure that all affected patients and families can benefit from the progress that has already occurred. The best is yet to come.

Acknowledgements
None.

Financial support and sponsorship

There was no financial support or sponsorship for the works.

Conflicts of interest
Casey Molecular Diagnostic Laboratory is a commercial genetic testing laboratory.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
- of special interest
- of outstanding interest

11. Strom SP, Goirin MB. Evaluation of autosomal dominant retinal dystrophy genes in an unaffected cohort suggests rare or private missense variants may often be benign. Mol Vis 2013; 19:980–985.

WES, with its wide applications in NGS testing, was billed as ‘panel killer’. Surprisingly, its efficacy has rarely been thoroughly analyzed. This study was the first task in the evaluation of WES and panel testing for retinal dystrophy.
13. Wang F, Wang H, Tuan H-F, et al. Next generation sequencing-based molecular diagnosis of retinitis pigmentosa: identification of a novel genotype-phenotype correlation and clinical refinements. Hum Genet 2014; 133:331–345. This was arguably one of the most significant works in the NGS of retinal dystrophy. CLN3, the gene many health care professionals specifically requested to exclude from NGS testing in the past, now became one of the most common nonsyndromic retinal dystrophy genes.


