Scaling New Heights in the Genetic Diagnosis of Inherited Retinal Dystrophies

Roser Gonzàlez-Duarte [1✉]

Email: rgonzalez@dbgen.com

Marta de Castro-Miró [2]

Miquel Tuson [2]

Valeria Ramírez-Castañeda [2]

Rebeca Valero Gils [2]

Gemma Marfany [1,2,3,4]

[1] Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[2] DBGen Ocular Genomics SL, Barcelona, Spain

[3] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Barcelona, Spain

[4] Institut de Biomedicina (IBUB-IRSJD), Universitat de Barcelona, Barcelona, Spain

Abstract

During the last 20 years, our group has focused on identifying the genes and mutations causative of inherited retinal dystrophies (IRDs). By applying massive sequencing approaches (NGS) in more than 500 familial and sporadic cases, we attained high diagnostic efficiency (85%) with a custom target gene panel and over 75% using whole exome sequencing (WES). Close to 40% of pathogenic alleles are novel mutations, which demand specific in silico tests and in vitro assays. Notably, missense variants are by far the most common type of mutation identified (around 40%), with small in-frame indels being less frequent (2%). To fill the gap of unsolved cases, when no candidate gene or only a single pathogenic allele has been identified, additional scientific and technical issues remain to be addressed. Reliable detection of genomic rearrangements and copy number variants (partial or complete), deep intronic mutations, variants that cause aberrant splicing events in retina-specific transcripts, functional assessment of hypomorphic missense alleles, mutations in regulatory sequences, the contribution of modifier genes to the IRD phenotype, and detection of low heteroplasmy mtDNA mutations are among the new challenges to be met.

Keywords

Inherited retinal dystrophies
Genetic diagnosis

Massive sequencing
Target gene sequencing
Whole exome sequencing
Precision medicine
IRD
NGS
WES
WGS

35.1. Introduction

Inherited retinal dystrophies (IRDs) show a prevalence of approximately 1:3000 individuals worldwide and are thus a relevant healthcare issue with a high impact both on patients and society. Identifying the causative gene and mutations of IRDs is desirable for several reasons: it secures the clinical diagnosis – particularly in difficult and syndromic cases – it paves the way for an accurate prognosis, it provides the basis for genetic counseling, it helps to better define the medical needs, and it is a prerequisite for precision medicine and patient selection for the upcoming gene and cell therapies (Marfany & Gonzàlez-Duarte 2015). NGS approaches have certainly revolutionized genetic diagnosis, and whole exome sequencing (WES) has become a basic methodological tool that also contributes to unveiling new causative IRD genes (a search in RetNet (https://sph.uth.edu/retnet/) reveals that more than 50 new genes have been reported in the last 4 years). Besides WES, targeted gene sequencing (TGS) provides a quick, reliable, and cost-effective genotype screening for routine genetic diagnosis. Although compared to WES the number of genes analyzed is much lower and new genes would remain undetected, with a rational design and a high depth of coverage, the mutation detection rate increases substantially.

A comparative survey of the success of TGS for IRDs in cohorts of more than 50 patients is presented in Table 35.1. Most groups screened from 105 to 283 genes, and the diagnosis yield ranged from 39.1% to 68%. In summary, differences in the positive diagnostic rates strongly depended on the coverage depth and the inclusion of relevant noncoding regions. The addition of newly reported genes, instead, did not have a linear impact on the global diagnostic yield but helped to solve clinically difficult cases. In our cohort, TGS using a custom panel of 332 IRD-related genes that covered all exons plus 65 intronic and noncoding regions reported to contain pathogenic variants resulted in a diagnosis yield close to 85% (Table 35.1, our data). Indeed, these high success rates, unimaginable 5 years ago, largely support the use of custom TGS in routine genetic diagnosis. WES remains the approach of choice in unsolved cases, when the contribution of novel genes might be envisaged, while whole genome sequencing (WGS) is becoming a feasible approach for detecting genome structural and intronic variants following cost reductions and the availability of user-friendly bioinformatics tools (Chen et al. 2017). To expand the number of solved cases, mtDNA can be easily added to nuclear DNA in the mutational screening, and we have added mtDNA sequencing to WES by adjusting DNA capture, depth of coverage, and bioinformatics analysis.

Table 35.1

Comparison of NGS custom gene panels for genetic diagnosis of IRDs

	Gene no.	Capture kit	NGS platform	Read depth	Patient no.	Diagnostic yield	Novel mutations

NR not reported

	Gene no.	Capture kit	NGS platform	Read depth	Patient no.	Diagnostic yield	Novel mutations
Ge et al. (2015)	195	NimbleGen (Roche)	Illumina HiSeq 2000	234×	105	49.5%	58.5%
Ellingford et al. (2016)	105	SureSelect (Agilent)	ABI SOLiD 5500 Illumina HiSeq	NR	537	50.5%	46.3%
Dockery et al. (2017)	254	NimbleGen (Roche)	Illumina MiSeq	NR	>750	68.3%	NR
Khan et al. 2017	105	SureSelect (Agilent)	SOLiD 4 (Life Technologies)	NR	115	39.1%	38%
Huang et al. (2018)	283	NimbleGen (Roche)	Illumina HiSeq 2000	400×	99	64.6%	76.6%
Our data by TGS	332	NimbleGen (Roche)	Illumina HiSeq 2000	>500×	73	84.9%	38.9%

NR not reported

35.2. Revisiting Mendelian Inheritance in IRDs

One of the most relevant incidental results of high-throughput genetic diagnoses is that clinically atypical presentations or patients with an ambiguous family history can also be successfully addressed (Jones et al. 2017). Since WES (and large IRD target panels with a comprehensive list of genes) includes all genes, data can be analyzed without previous genetic assumptions concerning Mendelian inheritance, resulting in higher genetic accuracy and even reclassification of patients. For instance, patients assumed to suffer from autosomal dominant retinitis pigmentosa (adRP) have been reclassified as presenting X-linked RP forms or even arRP in highly consanguineous families (Daiger et al. 2018). Some ocular disorders show overlapping clinical symptoms, and after target gene sequencing, the clinical entity can be accurately defined, e.g., simplex cases previously diagnosed as arRP or X-linked RP were reclassified as X-linked choroideremia (de Castro-Miró et al. 2018). This reclassification is crucial to opt for treatment, not only for gene and cell therapies but also for conventional diet- and drug-based treatments (e.g., Refsum disease; de Castro-Miró et al. 2016).

Also relevant for clinical management are the genetic diagnosis and reclassification of patients suffering from a syndromic IRD that initially went undetected. This is the case with mild mutations in ciliopathy-causative genes, whose first pathological trait is retinal degeneration, which will eventually involve other organs. Vice versa, cases initially considered rare syndromic disorders are accurately re-diagnosed as multi-Mendelian phenotypes after NGS (de Castro-Miró et al. 2018).

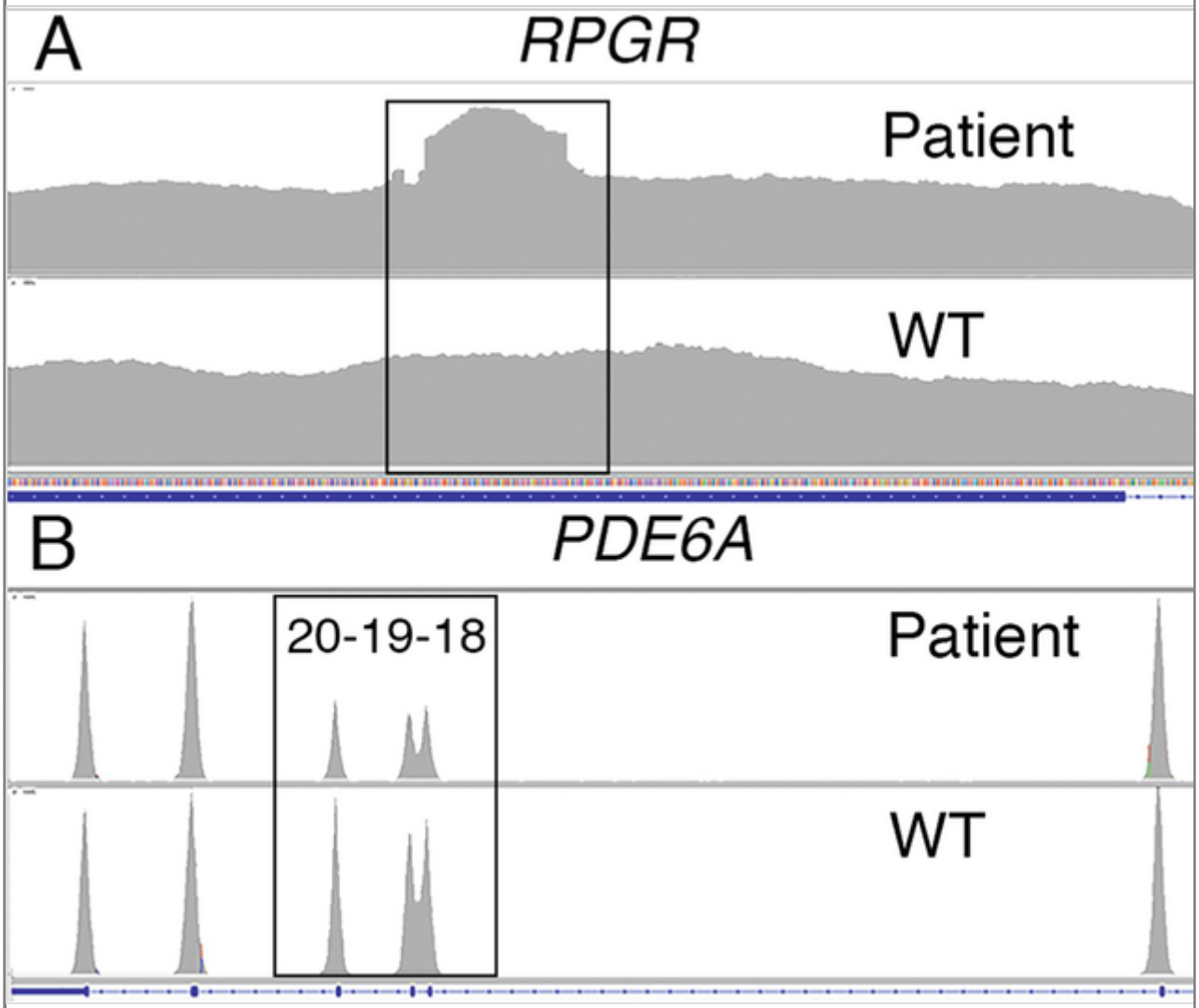
35.3. Delving into Difficult-to-Assess Genetic Variants

When using NGS, a high percentage of the pathogenic alleles identified in previously reported genes are novel mutations (from 38% to 60%; Table 35.1), which indicates that most pathogenic mutations are private to a few families. Classifying unreported variants that are nonsense or result in a frameshift (around 35% of the identified variants) as pathogenic and responsible for the

disease in patients is relatively straightforward. However, assigning pathogenicity to variants in the remaining cases is not a trivial task. How these issues are addressed underlies the difference in success rates and defines the quality and know-how of the genetic diagnosis labs. The detection of large indels within exons and the duplication/deletion of whole exons or genes (around 6% in our cohort) can be addressed if the sequencing coverage is high. In this context, TGS offers higher coverage and clearly overcomes WES. Two examples of successful identification of an internal exonic insertion (71 bp) and a heterozygous deletion of three contiguous exons (more than 5 kb), pathogenic mutations that would have otherwise remained undetected by WES, are shown in Fig. 35.1.

Fig. 35.1

High-coverage data in TGS enabled identification of (**a**) a hemizygous 71 bp duplication in the ORF15 exon of *RPGR* (c.2078_2148dup, p.Gln717Argfs∗4) and (**b**) a pathogenic deletion of exons 18–20 in *PDE6A*



In addition, among all identified mutations, around 45% are missense variants and 15% alter splicing events. The pathogenic effects of most missense and non-consensus-site splicing variants are difficult to assess and demand specific functional tests beyond in silico studies to support their

contribution to disease. Since a functional analysis is not always amenable, many of these variants end up in the omnium-gatherum of VUS (variants of unknown significance) category, thus accounting for a fraction of unsolved cases.

35.4. Casting Light upon Unsolved Cases

Pioneering work by several groups highlighted the relevant molecular effects of "hidden" mutations in introns and regulatory regions, which are usually not included in mutational screenings and are completely missed in clinical WES. However, reported deep intronic mutations in genes such as *CEP290*, *ABCA4*, *USH2A*, and others should be included in the diagnostic protocol of IRDs (Sangermano et al. 2017; den Hollander et al. 2006). Although TGS can include all reported intronic sequences, identifying novel variants and assigning pathogenicity become a Herculean task fraught with methodological and functional hurdles, particularly when the expression of the gene of interest is restricted to neurological or non-accessible tissues and in silico tests are merely indicative. The same obstacles hinder the identification of mutations in noncoding regulatory regions, which is clearly lagging behind, with very few successful cases (Radziwon et al. 2017; Small et al. 2016).

Valuable clues to undertaking a careful assessment of variants located in noncoding regions are either the identification of a single pathogenic allele that accords with the clinical entity or strong linkage to a particular candidate.

35.5. Increasing the Genetic Diagnosis Yield: Is There a Ceiling?

The number of causative genes of IRDs has certainly reached a plateau and will become a more and more sparse event. Screening methodologies and bioinformatics programs are evolving quickly into accurate and user-friendly tools to deal with "hidden" mutations in reported genes. Some intellectual challenges need be addressed to solve possibly rare and difficult cases: e.g., *cis* and *trans* mutations altering enhancer and regulatory regions, the impact of microexons (as reported in some neural pathologies (Porter et al. 2018) but whose mutational contribution to retinal disorders has yet to be assessed), and unreported retinal-specific isoforms (as happened with new mutations in *BBS8*, *RPGRIP1*, Jamshidi et al. 2018 Reference is now published in 2019. Therefore, it should read Jamishidi et al. 2019; Riazuddin et al. 2010).

Yet, there is most probably a "ceiling effect," and reaching a 100% success rate in genetic diagnosis of large cohorts might be unattainable. For obvious reasons, no biopsies of retinal tissue are available, and thus genetic shifts (as occurs in mtDNA heteroplasmy), somatic de novo mutations, and epigenetic changes (the effect of X-inactivation, somatic imprinting, and hyper– We wonder whether this hyphen should be smaller, like that after hypo- /hypo-methylation of regulatory sequences; Fahim & Daiger 2016) represent an unknown fraction of pathogenic mutations that are clearly beyond reach and may even underlie unilateral affectation.

35.6. Unveiling the Missing Link Between Rare and Common Diseases

ABCA4, the main gene of Stargardt's disease, has been leading breakthrough conceptual advances

in the identification of causative mutations, from illustrating allelic heterogeneity (mutations in the same gene cause distinct clinical retinal entities; Paloma et al. 2002) to reporting several deep intronic variants that cause aberrant splicing events (Sangermano et al. 2017) and identifying hypomorphic alleles as pathogenic when combined in trans with rare severe mutations (Zernant et al. 2018). If we consider that *ABCA4* hypomorphic alleles are not infrequent in the normal population, their association to disease blurs the once clear-cut boundary between rare Mendelian and common polygenic disorders. *ABCA4* variants may be either mutations for early-onset rare retinal dystrophies or – in homozygosity or compound heterozygosity with other hypomorphic alleles – risk variants for late-onset macular degeneration. The large amount of knowledge gathered on *ABCA4* mutations/variants enables scientists to show a continuum of genotype/phenotype correlations. Apparently, hypomorphic variants in *USH2A*, a gene involved in syndromic and non-syndromic retinal IRD, could have a similar role in age-related hypoacusia. The list of genes with both severe and hypomorphic alleles will surely increase in the near future.

35.7. Conclusion

Indeed, we envision an exciting future where TGS, WES, and WGS will be the leading actors in a new scenario, where modifier genes in rare diseases that account for incomplete penetrance and variable expressivity of the phenotype will be unveiled, and common hypomorphic variants that underlie susceptibility to late-onset and age-related disorders will be identified as the “missing link” between rare and common disorders.

References

AQ1

Chen Y, Zhao L, Wang Y et al (2017) SeqCNV: a novel method for identification of copy number variations in targeted next-generation sequencing data. BMC Bioinf 18:147

Daiger SP, Bowne SJ, Sullivan LS et al (2018) Molecular findings in families with an initial diagnose of autosomal dominant retinitis pigmentosa (adRP). Springer, Cham, pp 237–245

de Castro-Miró M, Tonda R, Escudero-Ferruz P et al (2016) Novel candidate genes and a wide spectrum of structural and point mutations responsible for inherited retinal dystrophies revealed by exome sequencing. PLoS One 11:1–19

de Castro-Miró M, Tonda R, Marfany G et al (2018) Novel mutation in the choroideremia gene and multi-Mendelian phenotypes in Spanish families. Br J Ophthalmol 102:1378–1386. https://doi.org/10.1136/bjophthalmol-2017-311427

den Hollander AI, Koenekoop RK, Yzer S et al (2006) Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. Am J Hum Genet 79:556–561

Dockery A, Stephenson K, Keegan D et al (2017) Target 5000: target capture sequencing for inherited retinal degenerations. Genes (Basel) 8:304

Ellingford JM, Barton S, Bhaskar S et al (2016) Molecular findings from 537 individuals with inherited retinal disease. J Med Genet jmedgenet:2016–103837

Fahim AT, Daiger SP (2016) The role of X-chromosome inactivation in retinal development and

disease. Adv Exp Med Biol 854:325–331

Ge Z, Bowles K, Goetz K et al (2015) NGS-based molecular diagnosis of 105 eyeGENE(®) probands with retinitis pigmentosa. Sci Rep 5:18287

Huang H, Chen Y, Chen H et al (2018) Systematic evaluation of a targeted gene capture sequencing panel for molecular diagnosis of retinitis pigmentosa. PLoS One 13:e0185237

Jones KD, Wheaton DK, Bowne SJ et al (2017) Next-generation sequencing to solve complex inherited retinal dystrophy: a case series of multiple genes contributing to disease in extended families. Mol Vis 23:470–481

Khan KN, Chana R, Ali N et al (2017) Advanced diagnostic genetic testing in inherited retinal disease: experience from a single tertiary referral centre in the UK National Health Service. Clin Genet 91:38–45

Marfany G, Gonzàlez-Duarte R (2015) Clinical applications of high-throughput genetic diagnosis in retinal dystrophies: present challenges and future directions. World J Med Genet 5:14–22

Paloma E, Coco R, Martínez-Mir A et al (2002) Analysis of *ABCA4* in mixed Spanish families segregating different retinal dystrophies. Hum Mutat 20:476–476

Porter RS, Jaamour F, Iwase S (2018) Neuron-specific alternative splicing of transcriptional machineries: implications for neurodevelopmental disorders. Mol Cell Neurosci 87:35–45

Radziwon A, Arno G, K Wheaton D et al (2017) Single-base substitutions in the CHM promoter as a cause of choroideremia. Hum Mutat 38:704–715

Riazuddin SA, Iqbal M, Wang Y et al (2010) A splice-site mutation in a retina-specific exon of BBS8 causes nonsyndromic retinitis pigmentosa. Am J Hum Genet 86:805–812

Sangermano R, Khan M, Cornelis SS et al (2017) ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. Genome Res 28:100–110

Small KW, DeLuca AP, Whitmore SS et al (2016) North Carolina macular dystrophy is caused by dysregulation of the retinal transcription factor PRDM13. Ophthalmology 123:9–18

Zernant J, Lee W, Nagasaki T et al (2018) Extremely hypomorphic and severe deep intronic variants in the *ABCA4* locus result in varying Stargardt disease phenotypes. Mol Case Stud 4:a002733