Published in final edited form as:

Novartis Found Symp. 2004 ; 255: 17–178.

Identifying retinal disease genes: how far have we come, how far do we have to go?

Stephen P. Daiger

Human Genetics Center, School of Public Health, and Department of Ophthalmology and Visual Science, The University of Texas Health Science Center, Houston, TX 77030, USA

Abstract

One of the great success stories in retinal disease (RD) research in the past decade has been identification of many of the genes and mutations causing inherited retinal degeneration. To date, more than 133 RD genes have been identified, encompassing many disorders such as retinitis pigmentosa, Leber congenital amaurosis, Usher syndrome and macular dystrophy. The most striking outcome of these findings is the exceptional heterogeneity involved: dozens of disease-causing mutations have been detected in most RD genes; mutations in many different genes can cause the same disease; and different mutations in the same gene may cause different diseases. Superimposed on this genetic heterogeneity is substantial clinical variability, even among family members with the same mutation. The RD genes involve many different pathways, and expression ranges from very limited (e.g. expressed in rod photoreceptors only) to ubiquitous. These findings raise several general questions in addition to the extraordinary number of specific, biological problems revealed. What fraction of the patient population can now be accounted for by the known RD genes? How many more RD genes will be found, and how should we find them? Are we dealing with just a handful of disease mechanisms or are there many different routes to retinal degeneration? How will this extreme heterogeneity affect our ability to diagnose and treat patients? These questions are considered in this summary.

During the past 15 years many of the genes and mutations causing inherited retinal diseases (RD genes) have been identified by a variety of methods. It is not a coincidence that this progress has paralleled completion of the human genome project: many of the technical advances in the genome project have enriched RD research. Also, identification of a complete set of human genes provides a powerful research tool in all areas of medicine, not just RD research. But RD research has been exceptionally productive, and might serve as a model for other types of inherited disease with both hopeful and disturbing implications.

The hopeful aspect is the demonstrated efficacy of what are now standard methods in medical genetics for gene identification, that is, ascertainment of patients and families, linkage mapping, positional candidate cloning and candidate gene screening. The disturbing aspect is just how complex retinal diseases have proven. The complexity includes:

- *Genetic heterogeneity*: mutations in different genes may cause the same retinal disease
- *Allelic heterogeneity*: many different disease-causing mutations are found in most RD genes
- *Phenotypic heterogeneity*: different mutations within the same gene may produce different clinical phenotypes and
- *Clinical heterogeneity*: the same mutation in different individuals, even within the same family, may produce different clinical consequences.

Whether these forms of heterogeneity, which are so characteristic of inherited retinal diseases, are true for other inherited conditions remains to be seen. (Inherited deafness may be as complicated; Petit et al 2001.) As distressing as the complexity may be for clinicians and patients, though, the fact that we know so much about the genes and mutations causing retinal diseases is a testament to the exceptional progress made in recent years. Nonetheless, there is a long row to hoe before we have a full appreciation of the molecular causes of inherited retinal diseases, especially among non-Western populations, and this information is only the first step in understanding pathogenic mechanisms and in providing treatments and cures.

Progress in identifying RD genes

There are many ways to measure progress in this field, for example, by the number of genes cloned, the number of mutations identified, or the number of investigators involved. Perhaps the most meaningful measure of progress is implied by one of the goals articulated in the Five Year Plan of the Foundation Fighting Blindness in 2000, ‘... to foster research leading to identification of the underlying genetic cause in 95% of patients with inherited retinal dystrophy ...’ (FFB Planning Document 2000).

That is, the question is not simply how many genes and mutations have been found but, rather, in what fraction of patients can a cause be identified? More explicitly, the issue is not ‘can mutations in this gene cause retinal disease’ but, instead, ‘do mutations in this gene cause disease in actual patients and, if so, in what fraction of patients’. This emphasis places the focus on prevalence, which is directly relevant to patient care, to the market for gene-specific treatments, and to the burden of these disorders on society.

In practice, disease-causing genes are usually identified without consideration of prevalence, say, by linkage mapping and cloning, and later screened in a large population of patients with appropriate phenotypes to establish prevalence. Therefore, one measure of progress in identifying RD genes is simply a count of known genes. The RetNet database maintains a list of cloned and/or mapped RD genes with information on clinical associations, the protein product, if known, and primary references and links to other sites (<http://www.sph.uth.tmc.edu/RetNet>). Figure 1 is a graph of the progress to date in mapping and cloning RD genes and Table 1 is a summary of the clinical categories for these genes, both derived from RetNet.

Figure 1 is based on the date of publication of the first article that reported chromosomal mapping of a RD locus and the first article that reported identification of the underlying disease-causing gene. In the case of RD genes identified by candidate gene screening, the ‘mapping’ and ‘cloning’ dates are the date of the first publication reporting identification of disease-causing mutations. Table 1 is simply a ‘head count’ of known RD genes. Please note that the clinical categories can be misleading because different mutations in the same gene may cause different diseases, yet the Table counts each gene once only. In cases where a gene falls into more than one category, it is counted in the category in which it was first described. For a list of all genes in all categories, see RetNet (<http://www.sph.uth.tmc.edu/RetNet/sum-dis.htm#B-diseases>).

Identification of RD genes progressed slowly from mapping the first X-linked retinitis pigmentosa (RP) gene in 1984 (Bhattacharya et al 1984) until the early 1990s, as seen in Fig. 1. Then, a number of technical developments, such linkage mapping using tandem repeat polymorphisms and improved software, greatly accelerated gene mapping. Within a few years thereafter the approaching completion of the human genome project provided tools to accelerate gene cloning. For most of the past 10 years, the number of mapped RD genes has increased roughly linearly and the number of cloned genes has increased more rapidly. From Fig. 1 there appears to be a flattening of both curves, around the time of publication of the first

draft of the human genome (by coincidence or otherwise). As noted below, since there are clearly many RD genes yet to be identified, the work is not yet done. A reasonable prediction is that laboratory techniques for mapping and identifying RD genes will improve markedly in the near future. However, the rate limiting steps are in ascertaining patients and families, and in clinical evaluation. Therefore, as always, the clinicians who work with affected individuals are the critical ingredient for continuing progress.

To date, 133 RD genes have been mapped and 87 have been cloned, as seen in Table 1. The Table counts genes causing many different retinal diseases, but even within a limited category such as retinitis pigmentosa (RP) there are numerous RD genes. For example, mutations in 12 genes cause autosomal dominant RP (ADRP), mutations in 15 others cause autosomal recessive RP (ARRP), and mutations in 5 cause X-linked RP (XLRP). In total, more than 49 genes are associated with RP, if syndromic forms are included. This illustrates the exceptional genetic heterogeneity of inherited retinal diseases.

Even these numbers are an underestimate, because they do not account for gene mutations implicated in more than one category of disease. For example, mutations in peripherin 2 (*RDS*) may cause either ADRP or autosomal dominant macular degeneration (Fellmar et al 1997), and mutations in usherin (*USH2A*) may cause either Usher syndrome or non-syndromic ARRP (Rivolta et al 2000).

Functional categories of RD genes

The proteins produced by the known RD genes fall into several functional categories, such as phototransduction, the visual cycle or retinal transcription factors (Phelan & Bok 2000, Rattner et al 1999). As might be expected, many RD genes are expressed exclusively or principally in photoreceptors and/or RPE cells, and play a role in 'obvious' visual functions. The limited number of relevant pathways, in comparison to the large number of RD genes, raises the hope that therapies may be focused on downstream components of these pathways rather than each individual RD gene. Therapies directed at slowing retinal cell apoptosis are an example of pathway targeting (Chader 2002).

However, several recently-identified RD genes are ubiquitously expressed, and do not fall comfortably into 'vision pathways'. Perhaps the most striking counter examples are three genes, *HPRP3*, *PRPF8* (RP13) and *PRPF31* (RP11). Mutations in these genes cause ADRP, and *only* ADRP (as far as is known), but each codes for a distinct, highly-conserved, essential protein component of the RNA splicing complexes found in all eukaryotes (Chakarova et al 2002, Mckie et al 2001, Vithana et al 2001). If nothing else, these findings demonstrate that our understanding of retinal development, cytology and biochemistry is more limited than generally recognized. They also illustrate the power of RD gene identification to reveal new aspects of retinal biology.

Progress in identifying disease-causing mutations in RD genes

Table 2 presents the number of unique disease-causing mutations reported for several selected RD genes. The data are based on pathogenic mutations listed in the Human Gene Mutation Database (<http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html>, Krawczak et al 2000). Additional resources for RD gene mutations are the Mutation Database of Retina International (<http://www.retina-international.org/sci-news/mutation.htm>) and individual RD gene entries in OMIM (<http://www3.ncbi.nlm.nih.gov/Omim/>, Hamosh et al 2002).

To date, rhodopsin mutations are the most numerous, with more than 100 reported, but other RD genes, such as *RDS* and *MYO7A*, have more than 50 reported mutations each. There is a correlation between the time since the first mutation was reported and the current total for each

gene—suggesting that the number of distinct mutations observed is related to the number of patients surveyed. Because of the large number of known mutations in some cases, there have been several attempts to find simplifying rules for the association between genotype and phenotype, with limited success so far. Some pathogenic mutations seem to cluster within protein domains (e.g. RP1; Berson et al 2001), whereas mutations in other genes are scattered throughout (e.g. rhodopsin; Hargrave 2001). In some cases loss-of-function mutations predominate, whereas, in others, missense mutations are more frequent (e.g. RP1 and rhodopsin, respectively). That is, global rules to explain the occurrence, molecular distribution and clinical consequences of RD gene mutations have not yet emerged.

The striking allelic heterogeneity of RD genes has disquieting implications. As noted below, some mutations in RD genes are ‘common’, relative to other mutations at the same locus, but the aggregate frequency of the rare or one-of-a-kind mutations is at least 50% in most cases. Thus to find all pathogenic mutations within a given gene it is necessary to scan the entire coding sequence, at least. This is a daunting task for large genes such as RP1 with 2156 amino acids or ABCA4 with 50 coding exons.

Of more concern, though, is that rare pathogenic mutations occur on a genetic background of numerous non-pathogenic variants, both polymorphic and rare. Simply because an amino acid substitution is found in a RD gene, and is rare, does not mean it is pathogenic. Worse, in positional candidate cloning of a mapped, dominant disease locus, simply because a rare missense mutation is tracking with disease does not prove pathogenicity. This is because rare, benign variants will be ‘common’, in aggregate, across the large non-recombinant regions usually implicated in positional cloning projects.

The extent of background variation is documented in a recent publication by Stephens et al (2001). The paper reports the numbers and types of sequence variants detected in sequencing 313 human genes in 82 individuals, representing 4 ethnic groups. A total of approximately 3900 polymorphic nucleotide substitutions (SNPs) were observed; 23% were variable within all groups, 25% were variable within one group only, and 38% were rare, i.e. found in one individual only. Of these, more than 50% lead to an amino acid substitution and roughly 1% introduce a premature stop codon. Put another way, in this study, on average

- an individual is heterozygous every 1.4 kb within coding sequences
- roughly half of these entail an amino acid substitution, and
- roughly 1 in 25 are rare (unique to an individual or family)

Thus rare, non-pathogenic amino acid substitutions will be found, inevitably, in screening projects in which the same gene is sequenced in many individuals, or in which several contiguous genes are sequenced in one individual. The conclusion: it is not sufficient to say that a mutation is pathogenic because it is rare—additional genetic and/or biochemical evidence must be adduced to make the case.

Prevalence of disease-causing mutations causing autosomal dominant retinitis pigmentosa

A misreading of Table 2 might suggest that all pathogenic mutations in RD genes are unique, one-of-a-kind events. This is not usually the case. In fact, of the scores of mutations reported at each locus, typically 2 or 3 are much more prevalent than others. Focusing specifically on ADRP, Table 3 lists pathogenic mutations which are found in multiple, unrelated families (Sohocki et al 2001, Bowne et al 1999, 2002). (Similar findings apply to other forms of inherited retinal disease.) ‘Unrelated’ in this context means that families sharing a mutation are usually not aware of each other’s existence but that, nonetheless, the mutation arose in a common

ancestor, perhaps hundreds of years ago. That is, where haplotype data are available, most prevalent mutations have been shown to arise by founder effects, not recurrent mutation.

For each gene in Table 3, a few mutations account for a large fraction of the total. For rhodopsin, the Pro23His mutation accounts for 40% of the total; for RP1 and IMPDH1, one mutation in each case, Arg677Ter and Asp226Asn, respectively, accounts for 50%. There is an important caveat to these observations: the prevalent mutations are always limited to a specific geographic group, consistent with historically recent founder events. For example, the rhodopsin Pro23His mutation is found in Americans of European origin but not Europeans (Farrar et al 1990). Thus any contributions to prevalence are limited to specific populations.

Progress in identifying mutations in patients with autosomal dominant retinitis pigmentosa

Taking adRP as a representative example of progress in identifying RD genes, it is now possible to identify a disease-causing mutation in 50–60% of patients. Table 4 shows the derivation of this estimate [references in RetNet]. As each ADRP gene was identified, the first reports were followed by screening — either by sequencing or mutation scanning (e.g. by SSCP or DGGE) of large collections of patients. Although the published studies are dissimilar in methodology, roughly comparable prevalences have been reported. Thus rhodopsin mutations are found in approximately 30% of adRP cases; mutations in four genes, PRPF31, RP1, RDS and IMPDH1, account for 5–8% each; and mutations in a few others account for 1–3% each. Table 4 lists the known ADRP genes in order of prevalence, with a cumulative prevalence of 60%.

This represents great progress toward the goal of identifying the cause in 95% of affected individuals, but the estimates come with *many* caveats. First, the selection of patients and screening methods differ between studies. Perhaps the greatest difference is in defining ‘autosomal dominant’ families. Second, each study is relatively small, and little or no attempt has been made to estimate confidence intervals. Third, all such studies have subtle biases in ascertainment of patients which may limit applicability to unselected populations.

Finally, the most important caveat is that, at best, the estimates are only applicable to the two major groups commonly studied: Americans of European origin and Europeans. Recent studies suggest that frequencies in Asia and elsewhere will be very different (Zhang et al 2002). This is explained, in part, by the predominance of founder mutations which are limited in geographic distribution. A reasonable prediction is that the genes and mutations identified in Americans and Europeans will be different from the most common causes in other populations. Further, of course, additional major genes will be identified in the ‘well-studied’ groups, too.

Implications and conclusions

By any reasonable measure, substantial progress has been made in identifying genes and mutations causing inherited retinal diseases. In spite of the exceptional heterogeneity observed, it is possible to detect disease-causing mutations in upwards of 60% of cases in certain well-defined patient populations, such as Americans with autosomal dominant retinitis pigmentosa.

Unfortunately, there is a very large gap between what can be done in theory and what is possible in practice. Detection of mutations in 60% of ADRP patients requires sequencing, or the equivalent, of six or more RD genes, totalling more than 60 000 kb of DNA over 40 exons. Any reasonable estimate of costs is in the thousands of dollars per patient (Table 5). Pre-screening for the most common mutations is much less expensive but will miss, perhaps, 50% of pathogenic variants. At present, routine mutation testing is not available for ADRP patients,

either commercially or in an academic setting. Developments in technology will eliminate this diagnostic bottleneck within a few years, we hope.

The era of gene mapping and positional ‘cloning’ is not over, notwithstanding the great progress to date. Whether only a few new RD genes will bring the total to 95% of patients, or many, is unknown. It is very likely that different sets of RD genes will predominate in other populations. That is, we have achieved much but we have a long way yet to go.

References

- Berson EL, Grimsby JL, Adams SM, et al. Clinical features and mutations in patients with dominant retinitis pigmentosa-1 (RP1). *Invest Ophthalmol Vis Sci* 2001;42:2217–2224. [PubMed: 11527933]
- Bhattacharya SS, Wright AF, Clayton JF, et al. Close genetic linkage between X-linked retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. *Nature* 1984;309:253–255. [PubMed: 6325945]
- Bowne SJ, Daiger SP, Hims MW, et al. Mutations in the RP1 gene causing autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 1999;8:2121–2128. [PubMed: 10484783]
- Bowne SJ, Sullivan LS, Blanton SH, et al. Mutations in the inosine monophosphate dehydrogenase 1 gene (MPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 2002;11:559–568. [PubMed: 11875050]
- Chader GJ. Animal models in research on retinal degenerations: past progress and future hope. *Vision Res* 2002;42:393–399. [PubMed: 11853755]
- Chakarova CF, Hims MM, Bolz H, et al. Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 2002;11:87–92. [PubMed: 11773002]
- Farrar GJ, Kenna P, Redmond R, et al. Autosomal dominant retinitis pigmentosa: absence of the rhodopsin proline-histidine substitution (codon 23) in pedigrees from Europe. *Am J Hum Genet* 1990;47:941–945. [PubMed: 2239971]
- Felbor U, Schilling H, Weber BHF. Adult vitelliform macular dystrophy is frequently associated with mutations in the peripherin/RDS gene. *Hum Mutat* 1997;10:301–309. [PubMed: 9338584]
- FFB Planning Document. Foundation Fighting Blindness. Owings Mills; MD, USA: 2000.
- Hamosh A, Scott AF, Amberger J, Bocchini C, Valle D, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2002;30:52–55. [PubMed: 11752252]
- Hargrave PA. Rhodopsin structure, function, and topography: the Friedenwald lecture. *Invest Ophthalmol Vis Sci* 2001;42:3–9. [PubMed: 11133841]
- Krawczak M, Ball EV, Fenton I, et al. Human gene mutation database — a biomedical information and research resource. *Hum Mutat* 2000;15:45–51. [PubMed: 10612821]
- McKie AB, McHale JC, Keen TJ, et al. Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum Mol Genet* 2001;10:1555–1562. [PubMed: 11468273]
- Petit C, Levilliers J, Hardelin JP. Molecular genetics of hearing loss. *Annu Rev Genet* 2001;35:589–646. [PubMed: 11700295]
- Phelan JK, Bok D. A brief review of retinitis pigmentosa and the identified retinitis pigmentosa genes. *Mol Vis* 2000;6:116–124. [PubMed: 10889272]
- Rattner A, Sun H, Nathans J. Molecular genetics of human retinal disease. *Annu Rev Genet* 1999;33:89–131. [PubMed: 10690405]
- Rivolta C, Sweklo EA, Berson EL, Dryja TP. Missense mutation in the USH2A gene: association with recessive retinitis pigmentosa without hearing loss. *Am J Hum Genet* 2000;66:1975–1978. [PubMed: 10775529]
- Sohocki MM, Daiger SP, Bowne SJ, et al. Prevalence of mutations causing retinitis pigmentosa and other inherited retinopathies. *Hum Mutat* 2001;17:42–51. [PubMed: 11139241]
- Stephens JC, Schneider JA, Tanguay DA, et al. Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 2001;293:489–493. [PubMed: 11452081]

- Vithana EN, Abu-Safieh L, Allen MJ, et al. A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell* 2001;8:375–381. [PubMed: 11545739]
- Zhang X, Fu W, Pang CP, Yeung KY. Screening for point mutations in rhodopsin gene among one hundred Chinese patients with retinitis pigmentosa. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2002;19:463–466. [PubMed: 12476415]

FIG. 1.
Mapped and cloned retinal diseases 1980–2002.

TABLE 1

Summary of selected retinal disease genes

Category	Mapped only	Cloned	Total
Total	46	87	133
Autosomal dominant retinitis pigmentosa	1	11	12
Autosomal recessive retinitis pigmentosa	5	10	15
X-linked retinitis pigmentosa	3	2	5
Usher syndrome	5	6	11
Bardet-Biedl syndrome	2	4	6
Leber congenital amaurosis	2	4	6
Autosomal dominant macular degeneration	4	4	8

TABLE 2

Summary of selected retinal disease mutations

Category	Gene symbol	Mutations reported
Autosomal dominant retinitis pigmentosa	RHO	103
	RDS	69
	RP1	14
	IMPDH1	9
Autosomal recessive Leber congenital amaurosis	RPE65	39
	GUCY2D	32
	AIP1	14
X-linked retinitis pigmentosa	RPGR	62
	RP2	30
Usher syndrome	MYO7A	56
	USH2A	31
	USH1C	7

TABLE 3

Common retinal disease mutations

Gene symbol	Mutation	% of total per gene
<i>RHO</i>	Pro23His	40
	Arg135Trp	3
	~5 others	6
<i>RDS</i>	Pro210Aig	25
	IVS2A>T	12
<i>RPI</i>	Arg677ter	50
	Leu726dd5	16
	Gly723ter	16
<i>IMPDH1(RP10)</i>	Asp226Asn	50
	Gly324Asp	25

TABLE 4

Prevalence of identifiable mutations in patients with autosomal dominant retinitis pigmentosa

Gene symbol	% of ADRP	Commutative %
<i>RHO</i>	30	30
<i>PRPF31</i> (<i>RP11</i>)	8	38
<i>RP1</i>	6	44
<i>RDS</i>	6	50
<i>IMPDH1</i> (<i>RP10</i>)	5	55
<i>HPRP3</i> (<i>RP18</i>)	3	58
<i>NRL</i>	1	59
others	<1	60

TABLE 5

Estimate of costs for mutation screening of patients with autosomal dominant retinitis pigmentosa

Item	Cost (SUS)	Commutative costs
Set patient file	100	100
Prepare and store DNAs	200	300
Screen for 'common' mutations by DHPLC	500	800
Sequence genes causing > 5% of ADRP	2000	2800
DHPLC and sequence remaining genes causing > 1%	3500	5300