Current Gene Discovery Strategies for Ocular Conditions

Priya Duggal^{*,1}, Grace Ibay², and Alison P. Klein^{*,1,2,3}

Many eye diseases are complex traits influenced by both genetic and environmental factors. Even for the common ocular conditions, such as refractive errors and glaucoma, there is a wide spectrum in the relative contribution of genetic and nongenetic factors in the development of disease. Some individuals develop disease, because they have inherited a single genetic mutation, whereas in others, the disease reflects the action of multiple genetic and/or environmental exposures. Recent advances in genetic technology have greatly enhanced our ability to identify the genetic variants underlying disease. In this review, we discuss how to determine whether an ocular phenotype has genetic components and if so, how to identify susceptibility genes associated with that phenotype. We also review the best approaches for identifying genetic variants of large effect (i.e., those with large relative risk) and genetic variants of smaller effect (i.e., those with small relative risk). Finally, we discuss how next-generation sequencing approaches will change the current paradigm for gene discovery. Figure 1 outlines the different genetic epidemiology approaches.

MENDELIAN AND COMPLEX INHERITANCE

For most ocular traits, there are both Mendelian and complex models of inheritance. Mendelian diseases result when a mutation in a single gene is sufficient to disrupt a biological pathway and lead to clinical disease. However, there could be several genes in a pathway that, when mutated, may cause the same disease phenotype. Mutations in these genes are often characterized as *dominant*, one mutant copy is sufficient to cause disease, or *recessive*, two mutant copies are necessary to cause disease. Mendelian diseases follow a predictable pattern of transmission in families that is often influenced by high (>80%) penetrance (the probability that an individual who has inherited a high-risk genotype develops disease). If the penetrance is complete or high, the Mendelian patterns of transmission are more evident. In Mendelian diseases, the actual genetic mutations may be specific to an extended family and very rare in the general population. Table 1 outlines examples of Mendelian eye disorders and the associated genes or loci.

Complex inheritance is caused by a combination of multiple genes, multiple environmental factors, or both genes and en-

vironmental factors. Genes underlying these complex diseases are more difficult to identify because there are multiple factors that contribute to the phenotype, and there can be a high degree of heterogeneity in the etiology of disease in any group of patients. The focus of most genetic research today is on identifying the factors that contribute to these complex diseases. Age-related macular degeneration (AMD) is a complex eye disorder in which genetic variation in the genes *CFH*¹⁻³ and *ARMS1* are known to increase progression of AMD, over and above the risk attributable to age and cigarette smoking. Identification of both the genetic and environmental causes of disease is critical to understanding the etiology of complex disease, especially since the penetrance of the underlying causal genes may be modified by environmental or lifestyle factors.

Although phenotypes are often considered discrete classifications based on disease status (i.e., presence or absence of cataract, myopic versus nonmyopic), many ocular diseases are a severe manifestation of an underlying quantitative phenotype. For example, cataract is a severe opacity of the lens, and myopia is a negative spherical equivalent requiring correction with an external lens for optimum visual acuity. Quantitative phenotypes often have complex inheritance patterns where several genes or environmental factors can act to alter the observed phenotype, and individuals who exceed a given threshold meet the clinical definition of "affected." Therefore, understanding the genetics of these underlying quantitative phenotypes can provide insight into the biology of common ocular diseases.

FAMILIAL AGGREGATION

If a disease is under genetic control (driven by genetic and not environmental factors), it will cluster in families, so that if disease is identified in one family member, there is a corresponding increased risk for other family members. To determine familial aggregation or clustering, data from cases and controls, selected cohorts of individuals, or the entire population can be used. If phenotypic information is available on patients and their parents, siblings or cousins, then a simple comparison of disease among the relatives compared with the general population is sufficient to show familial aggregation or an increased risk of disease among first-degree relatives (i.e., sibling and parent-child relationships). A *sibling* or *relative* recurrence risk ratio (λ_s) can be calculated that represents the risk of disease, given that an individual has an affected sibling, relative to the risk of disease in the overall population.⁴ Ideally, the disease status is collected through direct observation rather than only familial report, to lessen misclassification and reporting bias; however, this method requires examination of all family members, which is often difficult. If the status is indirectly reported through the index case/control or through population-based registries, then some information must be validated to define potential biases. In a cohort family, study of 269 pedigrees in the elderly Old Order Amish population, the sibling recurrence risk for different thresholds of myopia ranged from 2.36 (95% CI, 1.65–3.19) to 5.61 (95% CI, 3.06–

From the ¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health and the Departments of ²Oncology and ³Pathology, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, Maryland.

Supported by National Institutes of Health, National Eye Institute Grant EY017237 (APK).

Submitted for publication December 3, 2010; revised July 5 and August 4, 2011; accepted August 5, 2011.

Disclosure: P. Duggal, None; G. Ibay, None; A.P. Klein, None

*Each of the following is a corresponding author: Priya Duggal, Johns Hopkins University, Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205; pduggal@jhsph.edu.

Alison P. Klein, Sidney Kimmel Comprehensive Center, Johns Hopkins University, School of Medicine, 1550 Orleans Street, Baltimore, MD 21231; aklein1@jhmi.edu.

FIGURE 1. Overview of the genetic epidemiology approach.

9.34),⁵ suggesting a strong familial aggregation and possible genetic control for myopia.

In addition, individuals from an existing cohort, case-control, or cross-sectional study can be used to compare the risk of disease in first-degree relatives of cases and controls. Large population-based data, such as those from the Utah Population Database (UPDB), can also be used to determine familial aggregation. Luo et al.⁶ used the median Familial Standardized Incidence Ratio (FSIR) [the observed incidence of disease in a family compared with that expected in a standard population] to calculate an individual's familial risk of age-related maculopathy according to the occurrence of disease in the family. The median FSIR for this study was 3.95, indicating a strong familial aggregation of age-related maculopathy in Utah families.

For quantitative phenotypes, heritability can be estimated from familial correlations or variance components models where the observed phenotype variance and covariance among relatives can be partitioned into components reflecting

contribution of genes (G); household factors, or things shared within a household but not directly attributable to genetics (C); and residual environmental factors (E). The heritability (b^2) of a disease is the proportion of variance (V_p) attributable to additive genetic factors (V_g): $b^2 = V_g/V_p$. This ratio provides a measure of the importance of genetic factors in disease risk and can range between 0 and 1. However, estimates of heritability are population specific, and comparisons across populations must exercise caution. A recent review by Sanfilippo et al.⁷ summarizes heritability studies for various ocular disorders. The highest reported heritability was for small, hard drusen ($b^2 = 0.99$), suggesting that genetic factors may have strong control of the variation in this outcome. However, for other traits, there is a range of heritability from 0.15 to 0.91 for refractive error, 0.46 to 0.63 for corneal astigmatism, 0.29 to 0.67 for intraocular pressure, and 0.65 to 0.95 for central corneal thickness. These differences in heritability may reflect differences in methods, variability in the individual environmental effects, and differences between population-based studies and twin studies, the latter of which may have stricter control of age or possible cohort effects.

METHODS OF IDENTIFYING GENETIC VARIANTS

Once there is evidence supporting genetic contribution to a given disease, different methods can be used to identify the genes responsible for disease through some form of gene mapping. In general, these can be classified into two broad categories: linkage and association approaches. Genetic linkage is the violation of Mendel's law of independent assortment, which states that parental alleles at one gene are transmitted to the offspring independent of alleles at another gene. Genetic linkage will occur when two chromosomal loci are located physically close to one another on the same chromosome, so alleles at two loci co-segregate in a family across generations. Association is a more general statistical concept that tests for independence between genetic markers and observed phenotypes. Association is present when there is nonindependence between a genetic marker and a phenotype. Testing for association can be conducted in unrelated individuals (e.g., cases and controls) or among cases and their family members.

TABLE 1. Examples of Mendelian Forms of Eye Disease

Associated Disease	Mendelian Trait	Genes/Loci	Citation
Glaucoma	Rieger Syndrome	<i>RIEG1</i>	Murray et al. ³³
		<i>RIEG2</i>	Phillips et al. ³⁴
	Glaucoma with nail-patella syndrome	<i>LMX1B</i>	Vollrath et al. ³⁵
	Juvenile onset primary open angle glaucoma	<i>GLC1A (MYOC)</i>	Sheffield et al. ³⁶ ; Richards et al. ³⁷
	Primary congenital glaucoma	<i>CYP11B1 (GLC3A)</i>	Stoilov et al. ³⁸
Cataract		<i>GLC3B</i>	Akarsu et al. ³⁹
	Glaucoma with pigment dispersion syndrome	<i>GPDS1</i>	Andersen et al. ⁴⁰
	Congenital cataract	<i>CRYAA</i>	Pras et al. ⁴¹ ; Litt et al. ⁴²
	Cerulean type congenital cataract	<i>CCA1</i>	Armitage et al. ⁴³
	Congenital posterior polar cataract	<i>CRYAB</i>	Berry et al. ⁴⁴
	Anterior polar cataract	<i>CTAA2</i>	Berry et al. ⁴⁵
	X-linked myopia/Bornholm eye disease	<i>MYP1</i>	Schwartz et al. ⁴⁶
Myopia	Early or high myopia	<i>MYP2</i>	Young et al. ⁴⁷
	High myopia	<i>MYP3</i>	Young et al. ⁴⁸ ; Nurnberg et al. ⁴⁹
		<i>MYP12</i>	Paluru et al. ⁵⁰
		<i>MYP13</i>	Zhang et al. ⁵¹
		Decorin	Bredrup et al. ⁵²
Corneal Dystrophies	Congenital stromal dystrophy	<i>PIP5K3</i>	Jiao et al. ⁵³
	Francois-Neetens fleck (mouchetée) corneal dystrophy	<i>MHST6</i>	Vance et al. ⁵⁴
	Macular corneal dystrophy	<i>COL8A2</i>	Biswas et al. ⁵⁵
	Early onset Fuchs endothelial corneal dystrophy	<i>M1S1</i>	Tsujikawa et al. ⁵⁶ ; Ren et al. ⁵⁷
	Gelatinous drop-like corneal dystrophy		

Linkage Analysis

Linkage analysis uses the principle of genetic recombination to ask whether a disease phenotype is inherited jointly with a particular genetic marker in a given family, indicating that the disease locus is physically near the marker. If there is no linkage, the genetic marker is inherited independent of a disease phenotype; conversely, if linkage is present the genetic marker and disease are co-inherited. The strength of this co-inheritance is a function of how often there is chromosomal crossover—recombination—between the marker and disease loci. This function is one of physical distance between the marker and the disease loci. Thus, the closer the marker is to the disease loci, the stronger the evidence of linkage.

Linkage provides strong statistical evidence that a gene exists, although it may not identify the gene, but a chromosomal region or locus in which the putative gene resides. Linkage is a powerful and definitive genetic method of locating genes, but depends on having families with several affected or diseased members. Traditionally, this method has been used to identify most Mendelian disorders, although it can also be used for complex traits.

Recombination arises from an odd number of crossover events between any two loci. The recombination fraction (θ) is the probability that alleles at two genetic loci are inherited jointly in a family and is a function of the genetic distance between them, with the probability of recombination decreasing as the physical distance between the marker and disease loci decreases. Genes not linked will have 50% recombination of alleles during each meiosis—that is, the two alleles will be inherited jointly one half of the time due to chance alone. Loci that are linked have a recombination fraction $\theta < 0.5$, and loci that are in complete linkage show no recombination between them. The power to detect linkage depends on the genetic distance between the genetic marker and the true disease locus, with markers closer to the disease loci providing greater power. Genome-wide linkage studies enable the investigation of a putative susceptibility locus, or multiple loci, without any prior evidence of its position on the genome. Originally, these scans used microsatellite short tandem repeat polymorphic (STRP) markers that were evenly spaced, approximately 10 cM apart. Today, these scans are made using single nucleotide polymorphisms (SNPs) spaced approximately 1 cM apart. In addition, fine mapping or genotyping more markers in regions that have evidence suggestive of linkage will increase marker density and improve the power to localize a disease-causing gene. The power of linkage is based, in part, on the ability to accurately map meiotic events in a given family; thus, adding markers can allow for more accurate mapping of meiosis, thereby increasing power. The established criteria for assessing the statistical significance of linkage analysis in sibling pairs is that proposed by Lander and Kruglyak⁸ for declaring genome-wide significant ($P = 2.2 \times 10^{-5}$) and suggestive ($P = 7.4 \times 10^{-4}$) evidence of linkage, a well as of replication ($P = 0.01$). These thresholds control the genome-wide type 1 error rates while allowing for correlation between markers and are based on the assumption of an infinitely dense marker set.

The most powerful statistical model to test for linkage is the parametric or model-based linkage analysis. Parametric methods use both phenotypic and genotypic information from all family members and can be used with single or multiple markers, which may improve the information level, in a chromosome region. This method requires a specific genetic model for the disease, which includes specifying: (1) the number of genetic loci; (2) the mode of inheritance at each locus (dominant, recessive, codominant); (3) the frequency of the disease-causing allele; (4) the penetrance for each of its genotypes; (5) phenocopy rate, and (6) marker-allele frequencies. These pa-

rameters are used to compute a likelihood function that accounts for the parameters listed above and the unknown recombination fraction θ . The null hypothesis of no linkage (i.e., $H_0: \theta = 0.5$) is formally tested by comparing the likelihood of the null hypothesis to this same likelihood function maximized over the full range of θ —that is, the value of θ that best fits the observed family data. This test statistic is called the log-odds or LOD score (logarithm of the odds). Conventionally, a total LOD score >3.0 (odds 1000:1 in favor of linkage) is viewed as significant evidence of linkage when individual markers are considered.⁹ This corresponds to a P of approximately 0.0001.¹⁰ Parametric linkage analysis assumes the genetic model at the trait locus is known; however, many complex diseases do not have any established genetic model of inheritance. An alternative is to use nonparametric methods of linkage analysis that do not need a prespecified genetic model. However, nonparametric methods are less powerful than parametric methods, at least when the genetic model is correctly specified.

Nonparametric or model-free linkage analysis assesses the proportion of alleles shared identically by descent (IBD) among pairs or sets of relatives. IBD means the alleles have been transmitted to each individual in the relative pair from a common ancestor. The most commonly used relative pairs for qualitative trait analysis are affected sib pairs (ASP), although some methods may also use affected and unaffected relative pairs, and some include more distant relative pairs. In brief, the ASP statistic looks for IBD sharing in excess of what is expected by chance alone among affected sibling pairs. For traits with incomplete penetrance, the power to detect linkage may increase if the analysis is limited to affected pairs. However, the power of methods employing affected-only relatives may decrease if one of the cases in a pair is not due to inheritance of a gene (i.e., if one member is a phenocopy).^{9,11} The variance components approach can also serve as a form of model-free tests for linkage, as it can build on tests for excess allele sharing at marker loci for quantitative phenotypes.

Replication of results across independent linkage studies will confirm the existence of putative causal genes. However, negative results in subsequent linkage studies do not necessarily mean that the original results were a false-positive finding. Genetic heterogeneity, where several genes may be mutated and result in the same phenotype, may be at play—that is, the disease-causing genes in the first study are not the same genes as those involved in a second population, because of population differences or sampling variability. Another reason for nonreplication of linkage results is that the inherent nature of linkage studies has limited statistical power in samples of multiplex families to detect genes of modest effect. If a gene exerts only a small effect on risk in the families used for replication, there will not be enough power to confirm linkage.

Linkage analysis is a powerful tool for identifying loci involved in eye diseases, as evidenced by the numerous loci that have been identified for different traits. Table 2 outlines selected linkage results for eye diseases, showing success using different methods. In general, linkage analysis is best at identifying rare and less common variants of large effect and has not been as successful for complex disorders—perhaps because of the difficulty in identifying “genetic” families, since many complex disorders that are modified by environmental factors frequently have phenocopies or exhibit low penetrance. In addition, complex disorders are likely to have multiple loci involved in disease inheritance and the sample sizes necessary to identify these separate loci in linkage studies are difficult for a single study to collect.¹² However, linkage remains an important tool in statistical genetics that needs to

TABLE 2. Linkage Results for Select Ocular Traits

Trait	Linkage Method	Loci Identified	Citation
AMD	Parametric	1q25-q31	Klein et al. ⁵⁸
	Parametric and NPL	10q25; 9q33	Majewski et al. ⁵⁹
	NPL	1q31 and 10q26	Fisher et al. ⁶⁰
High-grade myopia	Parametric	18p11	Young et al. ⁴⁷
	Parametric	12q21	Young et al. ⁴⁸
	Parametric	17q21-q22	Paluru et al. ⁶¹
	Parametric	10q21.1	Nallasamy et al. ⁶²
	Parametric	5p15.33-p15.2	Lam et al. ⁶³
	NPL	7p15	Paget et al. ⁶⁴
	Parametric	4q22-q27	Zhang et al. ⁶⁵
	Parametric	2q37.1	Paluru et al. ⁵⁰
	Parametric	Xq23-25	Zhang et al. ⁵¹
	Parametric	22q12	Stambolian et al. ⁶⁶
Low-grade myopia	NPL	22q12	Klein et al. ⁶⁷
	NPL	11p13	Hammond et al. ⁶⁸
	NPL	3q26	Hammond et al. ⁶⁸
	NPL	4q12	Hammond et al. ⁶⁸
	NPL	8p23	Hammond et al. ⁶⁸
Refractive error	Parametric	1p36	Wojciechowski et al. ⁶⁹
	Parametric	5q22.1	Monemi et al. ⁷⁰
	Parametric	10p13-p14	Sarfarazi et al. ⁷¹ ; Rezaie et al. ⁷²
	Parametric	3q21-24	Wirtz et al. ⁷³ ; Kitsos et al. ⁷⁴
	Parametric	2p15-p16	Suriyapperuma et al. ⁷⁵
	NPL	15q11-q13	Allingham et al. ⁷⁶
	Parametric	8q23	Trifan et al. ⁷⁷
	Parametric	7q35-q36	Wirtz et al. ⁷⁸
	NPL	3p22-p21	Baird et al. ⁷⁹
	Parametric	2cen-q13	Stoilova et al. ⁸⁰
POAG	NPL	16p12-q13	Iyengar et al. ⁸¹

NPL (nonparametric linkage); AMD (age-related macular degeneration), QTL (quantitative trait loci).

be considered carefully for the ocular trait of interest. Strong linkage peaks, especially those that have been replicated in independent studies, provide robust evidence that a gene in that chromosomal region has a major effect on disease.

Association

Testing for genetic association involves comparing frequencies of marker alleles or genotypes between affected individuals and controls (either unrelated groups or sets of relatives). For a qualitative phenotype, the test hypothesis is complete independence between phenotype and allelic or genotypic classes. For a quantitative phenotype, the approach tests for a difference in the mean among different genotypes in a sample of unrelated individuals or whether the observed marker genotype explains a portion of the phenotypic variance/covariance among related individuals. Rejecting the null hypothesis of independence (or no effect of the marker on risk) shows there is a statistical association between marker and phenotype that may reflect a direct or indirect form of genetic control. The “direct effect” of a marker genotype implies causality, whereas an “indirect effect” means that the marker is linked to and in disequilibrium with some unobserved genetic variant controlling the phenotype. Even if a marker has no direct effect on the phenotype, an indirect relationship is useful for mapping unobserved causal genes. If alleles at two markers (either two variants or a marker and a causal variant) are associated in the population, they are said to be in disequilibrium. Linkage disequilibrium (LD) occurs when the two variants are tightly linked and do not recombine at every meiosis, even after several generations. If a new mutation arises next to a marker allele, it will be in complete LD with marker alleles immediately around it. In expanding populations, complete gametic association or complete LD gradually decays after many generations as the population approaches true Hardy-Weinberg equilibrium—a state of equilibrium between alleles and genotypes.

The tests of association performed in case-control studies can be applied to family-based designs, where the observed genotypes are treated as paired observations (case and control alleles or genotypes). The simplest family-based test of association uses the trio of parents and an affected child to compare alleles or genotypes transmitted to the affected child as the “case” and alleles not transmitted as the “control,”^{13,14} in an approach called the transmission/disequilibrium test (TDT).¹⁵ Excess transmission of a particular allele “case” is evidence of association. The TDT tests for linkage in the presence of disequilibrium—that is, the composite null hypothesis—is no linkage or no LD between the observed genetic marker and an unobserved causal gene. Rejecting the null hypothesis indicates evidence of both linkage and LD. Genotype data on both parents are needed for test validity, although only heterozygous parents actually contribute to the TDT statistic.

The most common association tests have used the candidate gene approach in which genetic variation in selected genes are tested for association with a disease or outcome. Traditionally, in the candidate gene approach, a set of markers within and around the coding sequence is identified and tested for association by using one of the methods described above. This process may include direct association of known functional variants or indirect association of genetic markers in LD with an unknown functional variant. These candidate gene studies have yielded some success. However, they are limited by our prior knowledge and our ability to suggest a biologically relevant gene.

In 2003, the International HapMap was established to aid in our ability to identify variation within the human genome and to determine patterns of LD.^{16,17} The International HapMap catalogs common variation in multiple geographic and ethnic populations and identifies haplotype blocks (regions in which there is correlation between genetic variants) and variants that “tag” these common haplotypes. The International HapMap

was used in the development of high-throughput, genome-wide SNP arrays, to offer the ability to assess most common genetic variation throughout the entire genome. These genome-wide association platforms capture much of the common variation in an individual's genome at a reasonable cost per sample. Unlike candidate gene studies, which are limited by a researcher's ability to identify a plausible candidate gene, the genome-wide association study (GWAS) approach provides a comprehensive approach for scanning the entire genome with good power to test common variants. To ensure a high rate of coverage for common variation in the human genome, these platforms include half a million to 5 million SNPs and rely on LD to capture regions of the genome without genotyping every known variant. The original GWAS platforms (<1 million SNPs) were designed to capture common genetic variation (>5%), and the newer platforms are more dense and include less common variants (1%–5%). Since these studies interrogate many SNPs, the thresholds for statistical significance are stringent (i.e., 10^{-7} – 10^{-8}), to limit type 1 errors, or the false-positive rate.¹⁸ In addition, because these studies are designed to identify common variants (variants present in at least 5%–10% of the population) with small to modest effect (relative risk in the range of 1.1–1.5) they require very large sample sizes (>1000 cases) to achieve good statistical power. Thus, despite the inclusion of less common variants on the newer platforms, GWAS is not powerful for the identification of rare variants unless the causal variants are in strong LD with a common variant.

Association studies of age-related macular degeneration led to the discovery of genetic variants strongly associated with AMD susceptibility. One of the most significant findings was a strong association between AMD and variants in and around complement factor H (*CFH*) on chromosome 1, region q32.^{1–3} Association studies in this region led to the discovery of additional complement genes showing strong association with AMD, including complement 2 (*C2*) and/or complement factor B (*CFB*)¹⁹ complement 3 (*C3*),²⁰ and complement factor I (*CFI*).²¹ Additional GWASs have identified several more genes that play a role in AMD susceptibility, including *TIMP3*, *SNY3*, and *LIPC*.^{22,23} The *CFH* finding was unique for GWASs with a small sample size ($n < 500$) and a large effect size (odds ratio range, 2–5). This is in contrast to the more typical GWAS of primary open-angle glaucoma requiring more than 3000 cases in the original and replicate studies to identify a modest effect size (odds ratio = 1.29) with *CAVI* and *CAV2*.²⁴ Similar GWASs have also yielded interesting results for myopia, high myopia, exfoliation glaucoma, and primary open-angle glaucoma (Table 3). However, identifying the causal variants within these identified genes is not always straightforward.

Replication of initial GWAS findings is a critical step in establishing that a genetic marker is truly associated with disease. As with linkage studies, several factors must be considered when assessing whether replication of a putative association is achieved. Standard criteria for replication of GWAS signals have been established by the NCI-NHGRI working groups.²⁵ In addition to standard epidemiologic criteria, such as ensuring that replication studies use the same definition for the phenotype, such studies should be well-powered and adhere to strict quality control standards. In addition, the genetic ancestry of each population should be considered, since allele frequencies are known to differ in continental populations and even within countries, and failure to replicate GWAS signals may be due to population differences.¹⁹ Conversely, additional studies in diverse populations can help narrow regions of association, by narrowing LD regions, as well as identifying additional genetic variants associated with disease to help explain differences in disease susceptibility across population groups.

Sequencing

Linkage analysis and GWAS are efficient at identifying genetic regions or loci that may harbor disease alleles for both large and small effect sizes, respectively. However, neither of these methods is designed to identify causal alleles, which is important for translational medicine, pharmacogenetics, and preventive screening. Sequencing provides the mechanism to discover all variants within a region which will include the causal variant(s). Traditional sequencing has used the Sanger method which is still the best for single gene regions. However, the development of next-generation sequencing technologies provides the ability to survey an individual genome at a base-pair resolution and the opportunity to identify many more genetic variants associated with disease on a finer scale for larger regions. Targeted resequencing of megabase regions of DNA, exome, and whole-genome sequencing using next-generation methods are rapidly emerging in genetic epidemiologic studies. These sequencing approaches allow for the rapid large-scale follow-up of regions identified in linkage studies or GWAS, as well as a powerful alternative to linkage studies for truly Mendelian diseases.

Exome sequencing focuses on the 1% of the human genome that encode proteins. This allows researchers to focus on genomic regions most likely to harbor deleterious variants. However, whole-genome sequencing has the advantage of complete coverage of the genome, including all regulatory regions that can have an important impact on human disease. This approach can be used in families with a Mendelian disease to identify rare genetic variants likely to be the cause of disease in that family. It is especially useful for highly penetrant diseases in which the novel rare variants can be tracked through the pedigree or in the identification of de novo germline mutations.²⁶ For example, a recent study suggested that mutations in the *ZNF644* gene were involved in high myopia.²⁷

In addition, sequencing can be used in population-based or family-based studies to identify rare causal variants for complex traits, especially after the gene has been localized through GWAS. The challenge of analyzing sequence data for complex diseases lies in correlating observed genetic variation with disease. Sequencing will identify many novel variants but determining which are associated with disease will require careful comparison to control individuals and development of statistical methods to evaluate multiple low allele frequencies. It is hypothesized that multiple rare variants in the same gene or different genes may act together in a polygenic model, each responsible for a very small risk.²⁸ The 1000 genomes project was established to catalog deep variation in the genome for alleles with frequencies 1% or greater in a large number of individuals.²⁹ These publicly available sequences will provide allele frequencies in the general population that can be used to compare to sequence in cases or those with an underlying trait. In addition, using linkage data from family-based studies to inform sequencing studies will help to narrow potential disease-associated variants. Furthermore, given the rapid decrease in sequencing costs, using exome sequencing to follow-up potential linkage regions may provide a more cost-efficient strategy than targeted resequencing for the identification of rare variants of large effect. Although genome sequencing will provide many new opportunities for gene discovery, to date, no large-scale genome sequencing studies aimed at disease gene discovery have been completed, largely because of the high-cost of genome sequencing; however, it is anticipated that the cost of sequencing will be ~\$1000 per individual in the coming year. Therefore, it is anticipated that many novel disease-associated genes will be identified in the coming decade by genome sequencing methods.

TABLE 3. Genome-Wide Association Results for Ocular Traits

Disease/Trait	Initial Sample Size	Replication Sample Size	Gene	Associated P-Value	Citation
AMD	96 white cases; 50 white controls	NR	<i>CFH</i>	4×10^{-8}	Klein et al. ²
AMD (wet)	96 Southeast Asian cases; 130 Southeast Asian controls	NR	<i>HTRA1</i>	8×10^{-12}	Dewan et al. ⁸²
AMD	979 European ancestry cases; 1,709 European ancestry controls	5,789 European ancestry cases; 4,234 European ancestry controls	<i>ARMS2</i> , <i>HTRA1</i> ; <i>CFH</i> ; <i>CFB</i> , <i>C2</i> ; <i>TIMP3</i> ; <i>CFE</i> ; <i>LIPC</i> ; <i>C3</i> ; <i>RREB1</i>	5×10^{-119} , 4×10^{-117} , 2×10^{-111} , 2×10^{-20} , 4×10^{-9} , 9×10^{-9} , 1×10^{-8} , 2×10^{-8} , 3×10^{-64} , 1×10^{-60} , 5×10^{-15} , 1×10^{-8}	Neale et al. ²³
AMD	293 family members; 391 white cases; 188 white controls	1,241 white cases; 622 white controls; 655 European ancestry cases; 1,244 European ancestry controls	<i>CFH</i> , <i>RMS2</i> , <i>HTRA1</i> ; <i>SKIV2L</i> , <i>BF</i> ; <i>C3</i> , <i>R102G</i>		Kopplin et al. ⁸³
Central corneal thickness	1,714 Australian twins and family; 1,759 UK twins and family; 249 Australian thin individuals; 251 thin individuals; 102 Australian individuals from extreme quantiles	NR	<i>ZNF469</i> ; <i>FOXO1</i>	9×10^{-11} , 5×10^{-10}	Lu et al. ⁸⁴
Central corneal thickness	1,445 European individuals	5,882 European individuals	<i>ZNF469</i> , <i>BANP</i> ; <i>AVGR8</i> ; <i>FOXO1</i> ; <i>PDE8A</i> ; <i>COL5A1</i>	6×10^{-22} , 4×10^{-9} , 1×10^{-8} , 1×10^{-8} , 5×10^{-8} , 1×10^{-18}	Vitart et al. ⁸⁵
Fuchs' corneal dystrophy	130 European ancestry cases; European ancestry; 260 controls	150 European ancestry cases; 150 European ancestry controls	<i>TCF4</i>		Baratz et al. ⁸⁶
Glaucoma, exfoliated	75 Icelandic cases; 14,474 Icelandic controls	254 European ancestry cases; 198 European ancestry controls	<i>LOXLI</i>	3×10^{-21}	Thorleifsson et al. ⁸⁷
POAG	305 Japanese cases; 355 Japanese controls	NR	<i>SRBD1</i>	3×10^{-9}	Meguro et al. ⁸⁸
POAG	1,263 Icelandic cases; 34,887 Icelandic controls	2,175 European ancestry cases; 2,064 European ancestry controls	<i>CAV1</i> , <i>CAV2</i>	2×10^{-11}	Thorleifsson et al. ²⁴
POAG	590 European ancestry cases; 3,956 European ancestry controls	892 European ancestry cases; 4,582 European ancestry controls	<i>CDKN2B-AS1</i> <i>TMC01</i>	1×10^{-14} ; 6×10^{-14}	Burdon et al. ⁸⁹
Myopia (pathologic)	297 Japanese cases; 934 Japanese controls	533 Japanese cases; 977 Japanese controls	<i>BLID</i> , <i>LOC399959</i>	2×10^{-7}	Nakanishi et al. ⁹⁰
Myopia (pathologic)	419 Han Chinese cases; 669 Han Chinese controls	843 Han Chinese cases and 1,960 Chinese cases; 2,525 Han Chinese controls and 3,117 Chinese controls	<i>MIPEP</i>	2×10^{-16}	Shi et al. ⁹¹
Myopia (pathologic)	102 Han Chinese cases; 335 Han Chinese controls	2,891 Han Chinese cases; 10,071 Han Chinese controls	<i>MYP11</i>	8×10^{-13}	Li et al. ⁹²
Vertical cup-disc ratio	7,360 European ancestry individuals	4,455 European ancestry individuals	<i>CDKN2B</i> ; <i>Six1</i> ; <i>SCYL1</i> ; <i>DCLK1</i> ; <i>CHEK2</i> ; <i>ATOH7</i> ; <i>BCAS3</i> ; <i>RERE</i> ; <i>ARID3A</i>	4×10^{-15} , 1×10^{-11} , 4×10^{-9} , 1×10^{-8} , 1×10^{-8} , 2×10^{-8} , 3×10^{-8} , 6×10^{-8} , 3×10^{-7}	Ramdas et al. ⁹³

(continues)

TABLE 3. (continued). Genome-Wide Association Results for Ocular Traits

Disease/Trait	Initial Sample Size	Replication Sample Size	Gene	Associated P-Value	Citation
Optic disc size (cup)	1,368 Australian twins; 848 United Kingdom individuals	NR	<i>ATOH1, RPTN1</i> ;	2×10^{-7} ; 2×10^{-7} ;	Macgregor et al. ⁹⁴
Optic disc size (disc)	1,368 Australian twins; 848 United Kingdom individuals	NR	<i>PBLD; ATOH7; MYPN; HSP90B3P; LRP1B; ZNF157</i> ;	2×10^{-10} ; 3×10^{-10} ;	Macgregor et al. ⁹⁴
Optic disc parameters	2,132 Indian ancestry individuals; 2,313 Malay ancestry individuals	9,326 European ancestry individuals	<i>CDC7, TGFB3; ATOH7; CARD10</i>	2×10^{-7} ; 3×10^{-7} ;	Khor et al. ⁹⁵
Optic disc parameters	7,360 individuals	4,455 individuals	<i>ATOH7, PBLD; CDC7, TGFB3</i>	8×10^{-17} ;	Ramdas et al. ⁹³
Refractive error	4,270 United Kingdom twins	13,414 European ancestry adults	<i>SALL1</i>	2×10^{-15}	Hysi et al. ⁹⁶
Refractive error	5,328 European ancestry individuals	10,280 European ancestry individuals	<i>RASGRF1</i>	2×10^{-9}	Solouki et al. ⁹⁷
			<i>GJD2, ACTC1, GOLG A8B</i>	2×10^{-14}	

Data for this table was generated from the NHGRI GWAS catalog available at the National Human Genome Research Institute, National Institutes of Health <http://www.genome.gov/gwastudies/> and reflects complete GWAS. NR (no replication sample); POAG (Primary Open Angle Glaucoma); AMD (Age related macular degeneration).⁹⁸

Additional Factors

All of these approaches hinge on well-characterized phenotypes (trait or disease), and strong measures of quality control. For both linkage and association studies these quality control measures can include determining whether there are Mendelian errors, laboratory or batch effects, genotyping errors, and properly called genotypes. In addition, we often restrict analyses to those with an allele frequency >1%, to ensure that there is enough power in the sample to detect a true finding. For population-based association studies, it is also important to consider confounding by ethnic ancestry or population stratification. Such confounding can occur if there are differences in allele frequencies between populations or subpopulations or admixtures that are all a part of the same study. In this situation, a spurious or masked association may result, owing to the underlying structure of the population and not the disease itself. Population stratification can be addressed using different methods, including genomic control, principal components and model-based clustering methods.³⁰⁻³² Family-based studies, by design, control for population stratification.

FOLLOW-UP STUDIES

This review focuses on methods of primary gene discovery. After the initial discovery that a gene or variant is associated with a disease, additional studies are needed to understand how both the normal and variant forms of the gene function. Gene expression (level of protein) is also influenced by genetic factors including variation in genetic sequence. Functional studies can be conducted to assess how a particular genetic variant alters gene expression. Conversely, linkage and association methods, including genome sequencing approaches can be used to identify genetic variation responsible for differences in expression levels. In this situation, the gene expression level is the phenotype or outcome of interest. In addition, studies of genes by environmental interaction are needed to determine the impact of environmental factors in conjunction with genetic factors on disease.

SUMMARY

Genetic factors play an important role in a variety of ocular disorders. The spectrum of the genetic contributions to disease range from rare high-penetrance genetic variants to common low-risk polymorphisms. GWAS allows us to identify common variants (frequency, >5%) that are associated with disease at a modest relative risk (<1.5). Yet, because these variations are frequent in the population, they may account for a substantive portion of the population risk. Family-based studies are best able to identify less frequent variants (<1%) that tend to have a stronger effect on disease (relative risk >3) and thereby can be used for risk assessment in families. Evolutionary history suggests rare mutations, like those found through linkage studies and by sequencing, are more likely to alter protein function and also more likely to have a large effect on disease. Natural selection has kept many of the mutations resulting in negative health effects at a lower frequency with the exception of diseases affecting people well after reproduction. These lower frequency variants are not detectable in most of the current GWASs, as these studies focus on effects of common variation (>5%) on risk. However, family-based linkage studies, as well as sequencing studies, should help identify rare high-penetrance variants. These primary approaches to gene discovery are complementary, and all approaches will be necessary to unravel the genetic basis of ocular disorders.

References

- Haines JL, Hauser MA, Schmidt S, et al. Complement factor H variant increases the risk of age-related macular degeneration. *Science*. 2005;308:419–421.
- Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308:385–389.
- Edwards AO, Ritter R 3rd, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science*. 2005;308:421–424.
- Risch N. Linkage strategies for genetically complex traits, II: the power of affected relative pairs. *Am J Hum Genet*. 1990;46:229–241.
- Peet JA, Cotch MF, Wojciechowski R, Bailey-Wilson JE, Stambolian D. Heritability and familial aggregation of refractive error in the Old Order Amish. *Invest Ophthalmol Vis Sci*. 2007;48:4002–4006.
- Luo L, Harmon J, Yang X, et al. Familial aggregation of age-related macular degeneration in the Utah population. *Vision Res*. 2008;48:494–500.
- Sanfilippo PG, Hewitt AW, Hammond CJ, Mackey DA. The heritability of ocular traits. *Surv Ophthalmol*. 2010;55:561–583.
- Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995;11:241–247.
- Ott J. *Handbook of Human Genetic Linkage*. Baltimore: Johns Hopkins University Press; 1999.
- Xu J, Meyers D, Pericak-Vance MA. Lod score analysis. In: Haines JL, Pericak-Vance MA, eds. *Approaches to Gene Mapping in Complex Human Diseases*. New York: Wiley-Liss; 1998:253–271.
- Bishop DT, Williamson JA. The power of identity-by-state methods for linkage analysis. *Am J Hum Genet*. 1990;46:254–265.
- Suarez BK, Hampe GL, VanEerdewegh P. Problems of replicating Linkage claims in psychiatry. In: Gershon ES, Cloninger CR, ed. *Genetic Approaches to Mental Disorders*. Washington, DC: American Psychological Association; 1994:23–46.
- Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet*. 1987;51:227–233.
- Rubinstein P, Ginsberg-Fellner F, Falk C. Genetics of type I diabetes mellitus: a single, recessive predisposition gene mapping between HLA-B and GLO with an appendix on the estimation of selection bias. *Am J Hum Genet*. 1981;33:865–882.
- Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet*. 1993;52:506–516.
- The International HapMap Project. *Nature*. 2003;426:789–796.
- A haplotype map of the human genome. *Nature*. 2005;437:1299–1320.
- Ziegler A, König I. *A Statistical Approach to Genetic Epidemiology*. Singapore: Wiley-VCH; 2010.
- Gold B, Merriam JE, Zernant J, et al. Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. *Nat Genet*. 2006;38:458–462.
- Yates JR, Sepp T, Matharu BK, et al. Complement C3 variant and the risk of age-related macular degeneration. *N Engl J Med*. 2007;357:553–561.
- Fagerness JA, Maller JB, Neale BM, Reynolds RC, Daly MJ, Seddon JM. Variation near complement factor I is associated with risk of advanced AMD. *Eur J Hum Genet*. 2009;17:100–104.
- Chen W, Stambolian D, Edwards AO, et al. Genetic variants near TIMP3 and high-density lipoprotein-associated loci influence susceptibility to age-related macular degeneration. *Proc Natl Acad Sci USA*. 2010;107:7401–7406.
- Neale BM, Fagerness J, Reynolds R, et al. Genome-wide association study of advanced age-related macular degeneration identifies a role of the hepatic lipase gene (LIPC). *Proc Natl Acad Sci USA*. 2010;107:7395–7400.
- Thorleifsson G, Walters GB, Hewitt AW, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. *Nat Genet*. 2010;42:906–909.
- Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature*. 2007;447:655–660.
- Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2010;42:30–35.
- Shi Y, Li Y, Zhang D, et al. Exome sequencing identifies ZNF644 mutations in high myopia. *PLoS Genet*. 2011;7:e1002084.
- Ramdas WD, Amin N, van Koolwijk LM, et al. Genetic architecture of open angle glaucoma and related determinants. *J Med Genet*. 2011;48:190–196.
- Durbin RM, Abecasis GR, Altshuler DL, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061–1073.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–959.
- Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55:997–1004.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38:904–909.
- Murray JC, Bennett SR, Kwitek AE, et al. Linkage of Rieger syndrome to the region of the epidermal growth factor gene on chromosome 4. *Nat Genet*. 1992;2:46–49.
- Phillips JC, del Bono EA, Haines JL, et al. A second locus for Rieger syndrome maps to chromosome 13q14. *Am J Hum Genet*. 1996;59:613–619.
- Vollrath D, Jaramillo-Babb VL, Clough MV, et al. Loss-of-function mutations in the LIM-homeodomain gene, LMX1B, in nail-patella syndrome. *Hum Mol Genet*. 1998;7:1091–1098.
- Sheffield VC, Stone EM, Alward WL, et al. Genetic linkage of familial open angle glaucoma to chromosome 1q21–q31. *Nat Genet*. 1993;4:47–50.
- Richards JE, Lichter PR, Boehnke M, et al. Mapping of a gene for autosomal dominant juvenile-onset open-angle glaucoma to chromosome 1q. *Am J Hum Genet*. 1994;54:62–70.
- Stoilov I, Akarsu AN, Sarfarazi M. Identification of three different truncating mutations in cytochrome P4501B1 (CYP1B1) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the GLC3A locus on chromosome 2p21. *Hum Mol Genet*. 1997;6:641–647.
- Akarsu AN, Turacli ME, Aktan SG, et al. A second locus (GLC3B) for primary congenital glaucoma (Buphthalmos) maps to the 1p36 region. *Hum Mol Genet*. 1996;5:1199–1203.
- Andersen JS, Pralea AM, DelBono EA, et al. A gene responsible for the pigment dispersion syndrome maps to chromosome 7q35–q36. *Arch Ophthalmol*. 1997;115:384–388.
- Pras E, Frydman M, Levy-Nissenbaum E, et al. A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. *Invest Ophthalmol Vis Sci*. 2000;41:3511–3515.
- Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, Weleber RG. Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. *Hum Mol Genet*. 1998;7:471–474.
- Armitage MM, Kivlin JD, Ferrell RE. A progressive early onset cataract gene maps to human chromosome 17q24. *Nat Genet*. 1995;9:37–40.
- Berry V, Francis P, Reddy MA, et al. Alpha-B crystallin gene (CRYAB) mutation causes dominant congenital posterior polar cataract in humans. *Am J Hum Genet*. 2001;69:1141–1145.
- Berry V, Ionides AC, Moore AT, Plant C, Bhattacharya SS, Shiels A. A locus for autosomal dominant anterior polar cataract on chromosome 17p. *Hum Mol Genet*. 1996;5:415–419.
- Schwartz M, Haim M, Skarsholm D. X-linked myopia: Bornholm eye disease: linkage to DNA markers on the distal part of Xq. *Clin Genet*. 1990;38:281–286.
- Young TL, Ronan SM, Drahozal LA, et al. Evidence that a locus for familial high myopia maps to chromosome 18p. *Am J Hum Genet*. 1998;63:109–119.
- Young TL, Ronan SM, Alvear AB, et al. A second locus for familial high myopia maps to chromosome 12q. *Am J Hum Genet*. 1998;63:1419–1424.
- Nurnberg G, Jacobi FK, Broghammer M, et al. Refinement of the MYP3 locus on human chromosome 12 in a German family with

- Mendelian autosomal dominant high-grade myopia by SNP array mapping. *Int J Mol Med*. 2008;21:429–438.
50. Paluru PC, Nallasamy S, Devoto M, Rappaport EF, Young TL. Identification of a novel locus on 2q for autosomal dominant high-grade myopia. *Invest Ophthalmol Vis Sci*. 2005;46:2300–2307.
  51. Zhang Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF. Novel locus for X linked recessive high myopia maps to Xq23–q25 but outside MYP1. *J Med Genet*. 2006;43:e20.
  52. Bredrup C, Knappskog PM, Majewski J, Rodahl E, Boman H. Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. *Invest Ophthalmol Vis Sci*. 2005;46:420–426.
  53. Jiao X, Munier FL, Schorderet DF, et al. Genetic linkage of Francois-Neetens fleck (mouchetee) corneal dystrophy to chromosome 2q35. *Hum Genet*. 2003;112:593–599.
  54. Vance JM, Jonasson F, Lennon F, et al. Linkage of a gene for macular corneal dystrophy to chromosome 16. *Am J Hum Genet*. 1996;58:757–762.
  55. Biswas S, Munier FL, Yardley J, et al. Missense mutations in COL8A2, the gene encoding the alpha2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy. *Hum Mol Genet*. 2001;10:2415–2423.
  56. Tsujikawa M, Kurahashi H, Tanaka T, et al. Identification of the gene responsible for gelatinous drop-like corneal dystrophy. *Nat Genet*. 1999;21:420–423.
  57. Ren Z, Lin PY, Klintworth GK, et al. Allelic and locus heterogeneity in autosomal recessive gelatinous drop-like corneal dystrophy. *Hum Genet*. 2002;110:568–577.
  58. Klein ML, Schultz DW, Edwards A, et al. Age-related macular degeneration. Clinical features in a large family and linkage to chromosome 1q. *Arch Ophthalmol*. 1998;116:1082–1088.
  59. Majewski J, Schultz DW, Weleber RG, et al. Age-related macular degeneration—a genome scan in extended families. *Am J Hum Genet*. 2003;73:540–550.
  60. Fisher SA, Abecasis GR, Yashar BM, et al. Meta-analysis of genome scans of age-related macular degeneration. *Hum Mol Genet*. 2005;14:2257–2264.
  61. Paluru P, Ronan SM, Heon E, et al. New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol Vis Sci*. 2003;44:1830–1836.
  62. Nallasamy S, Paluru PC, Devoto M, Wasserman NF, Zhou J, Young TL. Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota. *Mol Vis*. 2007;13:229–236.
  63. Lam CY, Tam PO, Fan DS, et al. A genome-wide scan maps a novel high myopia locus to 5p15. *Invest Ophthalmol Vis Sci*. 2008;49:3768–3778.
  64. Paget S, Julia S, Vitezica ZG, Soler V, Malecaze F, Calvas P. Linkage analysis of high myopia susceptibility locus in 26 families. *Mol Vis*. 2008;14:2566–2574.
  65. Zhang Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF. A new locus for autosomal dominant high myopia maps to 4q22–q27 between D4S1578 and D4S1612. *Mol Vis*. 2005;11:554–560.
  66. Stambolian D, Ibay G, Reider L, et al. Genomewide linkage scan for myopia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 22q12. *Am J Hum Genet*. 2004;75:448–459.
  67. Klein AP, Duggal P, Lee KE, Klein R, Bailey-Wilson JE, Klein BE. Confirmation of linkage to ocular refraction on chromosome 22q and identification of a novel linkage region on 1q. *Arch Ophthalmol*. 2007;125:80–85.
  68. Hammond CJ, Andrew T, Mak YT, Spector TD. A susceptibility locus for myopia in the normal population is linked to the PAX6 gene region on chromosome 11: a genomewide scan of dizygotic twins. *Am J Hum Genet*. 2004;75:294–304.
  69. Wojciechowski R, Moy C, Ciner E, et al. Genomewide scan in Ashkenazi Jewish families demonstrates evidence of linkage of ocular refraction to a QTL on chromosome 1p36. *Hum Genet*. 2006;119:389–399.
  70. Monemi S, Spaeth G, Dasilva A, et al. Identification of a novel adult-onset primary open angle glaucoma (POAG) gene on 5q22.1. *Hum Mol Genet*. 2005;14(6):725–733.
  71. Sarfarazi M, Child A, Stoilova D, et al. Localization of the fourth locus (GLC1E) for adult-onset primary open-angle glaucoma to the 10p15–p14 region. *Am J Hum Genet*. 1998;62:641–652.
  72. Rezaie T, Child A, Hitchings R, et al. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science*. 2002;295:1077–1079.
  73. Wirtz MK, Samples JR, Kramer PL, et al. Mapping a gene for adult-onset primary open-angle glaucoma to chromosome 3q. *Am J Hum Genet*. 1997;60:296–304.
  74. Kitsos G, Eiberg H, Economou-Petersen E, et al. Genetic linkage of autosomal dominant primary open angle glaucoma to chromosome 3q in a Greek pedigree. *Eur J Hum Genet*. 2001;9:452–457.
  75. Suriyapperuma SP, Child A, Desai T, et al. A new locus (GLC1H) for adult-onset primary open-angle glaucoma maps to the 2p15–p16 region. *Arch Ophthalmol*. 2007;125:86–92.
  76. Allingham RR, Wiggs JL, Hauser ER, et al. Early adult-onset POAG linked to 15q11–13 using ordered subset analysis. *Invest Ophthalmol Vis Sci*. 2005;46:2002–2005.
  77. Trifan OC, Traboulsi EI, Stoilova D, et al. A third locus (GLC1D) for adult-onset primary open-angle glaucoma maps to the 8q23 region. *Am J Ophthalmol*. 1998;126:17–28.
  78. Wirtz MK, Samples JR, Rust K, et al. GLC1F, a new primary open-angle glaucoma locus, maps to 7q35–q36. *Arch Ophthalmol*. 1999;117:237–241.
  79. Baird PN, Foote SJ, Mackey DA, Craig J, Speed TP, Bureau A. Evidence for a novel glaucoma locus at chromosome 3p21–22. *Hum Genet*. 2005;117:249–257.
  80. Stoilova D, Child A, Trifan OC, Crick RP, Coakes RL, Sarfarazi M. Localization of a locus (GLC1B) for adult-onset primary open angle glaucoma to the 2cen-q13 region. *Genomics*. 1996;36:142–150.
  81. Iyengar SK, Klein BE, Klein R, et al. Identification of a major locus for age-related cortical cataract on chromosome 6p12–q12 in the Beaver Dam Eye Study. *Proc Natl Acad Sci USA*. 2004;101:14485–14490.
  82. Dewan A, Liu M, Hartman S, et al. HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science*. 2006;314:989–992.
  83. Kopplin LJ, Igo RP Jr, Wang Y, et al. Genome-wide association identifies SKIV2L and MYRIP as protective factors for age-related macular degeneration. *Genes Immun*. 2010;11(8):609–621.
  84. Lu Y, Dimasi DP, Hysi PG, et al. Common genetic variants near the Brittle cornea syndrome locus ZNF469 influence the blinding disease risk factor central corneal thickness. *PLoS Genet*. 2010;6:e1000947.
  85. Vitart V, Bencic G, Hayward C, et al. New loci associated with central cornea thickness include COL5A1, AKAP13 and AVGR8. *Hum Mol Genet*. 2010;19:4304–4311.
  86. Baratz KH, Tosakulwong N, Ryu E, et al. E2–2 protein and Fuchs's corneal dystrophy. *N Engl J Med*. 2010;363:1016–1024.
  87. Thorleifsson G, Magnusson KP, Sulem P, et al. Common sequence variants in the LOXL1 gene confer susceptibility to exfoliation glaucoma. *Science*. 2007;317:1397–1400.
  88. Meguro A, Inoko H, Ota M, Mizuki N, Bahram S. Genome-wide association study of normal tension glaucoma: common variants in SRBD1 and ELOVL5 contribute to disease susceptibility. *Ophthalmology*. 2010;117:1331–1338 e1335.
  89. Burdon KP, Macgregor S, Hewitt AW, et al. Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. *Nat Genet*. 2011;43:574–578.
  90. Nakanishi H, Yamada R, Gotoh N, et al. A genome-wide association analysis identified a novel susceptibility locus for pathological myopia at 11q24.1. *PLoS Genet*. 2009;5:e1000660.
  91. Shi Y, Qu J, Zhang D, et al. Genetic variants at 13q12.12 are associated with high myopia in the han chinese population. *Am J Hum Genet*. 2011;88:805–813.
  92. Li Z, Qu J, Xu X, et al. A genome-wide association study reveals association between common variants in an intergenic region of 4q25 and high-grade myopia in the Chinese Han population. *Hum Mol Genet*. 2011;20:2861–2868.

93. Ramdas WD, van Koolwijk LM, Ikram MK, et al. A genome-wide association study of optic disc parameters. *PLoS Genet.* 2010;6:e1000978.
94. Macgregor S, Hewitt AW, Hysi PG, et al. Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. *Hum Mol Genet.* 2010;19:2716-2724.
95. Khor CC, Ramdas WD, Vithana EN, et al. Genome-wide association studies in Asians confirm the involvement of ATOH7 and TGFB3, and further identify CARD10 as a novel locus influencing optic disc area. *Hum Mol Genet.* 2011;20:1864-1872.
96. Hysi PG, Young TL, Mackey DA, et al. A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. *Nat Genet.* 2010;42:902-905.
97. Solouki AM, Verhoeven VJ, van Duijn CM, et al. A genome-wide association study identifies a susceptibility locus for refractive errors and myopia at 15q14. *Nat Genet.* 2010;42:897-901.
98. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA.* 2009;106:9362-9367.