The ABCA4 Gene and Age-Related Macular Degeneration

Innocence or Guilt by Association

The controversy regarding the potential role of the ABCA4 gene (formerly ABCR) and age-related macular degeneration (ARMD) began in 1997, when Allikmets et al1 proposed that this gene might account for 16% of ARMD cases based on the initial association study of a subset of ABCA4 alleles. This gene, which has been convincingly shown to be responsible for most cases of autosomal recessive Stargardt disease and more recently for a subset of patients with cone-rod dystrophy and retinitis pigmentosa,2,3 encodes a transmembrane protein that seems to use adenosine triphosphate to actively transport all-trans retinal out of the outer segments of the rod and cone photoreceptors.4,5 Since that initial report, there has been a flurry of studies in which some have attempted to confirm the association of ABCA4 with ARMD, while others have not found evidence of such a relationship.6,9

See also page 745

The article by Guymer et al10 in this issue of the Archives highlights some of the crucial issues that challenge not only those who are studying the genetic bases of ARMD, but all investigators who are studying the genetics of common and complex disorders. At the core of this debate is how we determine the validity and biological relevance of an association between a particular gene and a disease. A variation in a gene, even when it is in the coding region and causes a change in the amino acid sequence of the protein product, does not necessarily lead to a disease-causing state. Even if one can make a convincing argument that the structure and/or function of the protein is altered or even disrupted, this might not necessarily cause a particular disease if there are other proteins that can provide compensatory function. At the same time, one must be cautious in discounting a valid association between mutations in a gene and a disease because of incomplete penetrance. The family studies conducted with grandparents of patients with Stargardt disease have been criticized for ascertainment bias and/or inadequate statistical power.7,11 However, studies that claim that the presence of a particular ABCA4 variant in a family member who lacks any evidence of ARMD disproves a disease association are also limited by a lack of statistical power and an inability to estimate the extent that incomplete penetrance may play a role in ARMD.8 Guymer and colleagues illustrate this issue clearly in their reporting of the G1961E allele in unaffected family members with ARMD, and they accurately present their findings within the correct genetic context that considers incomplete penetrance. The observance of such cases does not exclude the potential for an ABCA4 allele to contribute to ARMD, but it does place major demands on investigators who wish to use family studies to argue in support of an association. Family linkage or association studies that seek to demonstrate a statistically valid association between a specific gene allele and ARMD must have sufficiently large families and/or large numbers of families to present a convincing argument, especially if there is a high degree of incomplete penetrance and/or genetic heterogeneity. For ARMD, this is especially challenging since we suspect that there is considerable genetic heterogeneity, and large ARMD families are relatively rare because of the late onset of the disease. When one performs linkage analysis with highly polymorphic markers that flank a candidate gene, one essentially collapses all of the potentially disease-causing mutations into a single analysis, thus increasing the power to detect linkage. When we employed this strategy to investigate the ABCA4 locus and ARMD, no evidence of linkage was found.12 This does not exclude the potential of this gene to contribute to ARMD, but the negative results do argue that the contribution (collectively of all of the ABCA4 variants) to ARMD must be small compared with those loci for which there is evidence of linkage.

There are 2 additional issues that potentially confound the current association studies of ABCA4 and ARMD. One, which is addressed in the article by Guymer and colleagues, is the fact that different populations may show different frequencies of specific gene alleles. Even in this study, the control population has a different age distribution than the ARMD population and thus is incompletely matched. It is absolutely crucial for association studies to have a genetically matched control population with respect to the disease-affected group. The high prevalence of the G1961E allele in the healthy individuals of Somali ancestry and the range of G1961E allele frequencies observed between the US and Swiss patients illustrate the need to constantly consider this problem. Recently, some investigators suggested that one should perform several genotypings of unlinked markers that are in equilibrium with respect to the disease within both the control and affected populations to compare allele frequencies and determine if the populations are comparable on a molecular basis.

The high prevalence of the G1961E allele in the healthy Somali population tends to minimize the argument that this allele is disease causing because there is no reported high incidence of either ARMD or Stargardt disease.
disease in that population. Though less likely, it is theoretically possible that in a different ethnic group, the allele could have a much more detrimental effect depending on the prevalence of modifier gene alleles, environmental factors, or the frequency of other ABCA4 alleles within the population. There are a number of animal models of genetic mutations that cause significantly different phenotypes, disease severity, sex differences, and varying degrees of penetrance when transferred to different genetic backgrounds or in response to behavioral and environmental influences.13,14

Finally, there is the issue of phenocopies and the potential of miscategorizing end-stage Stargardt disease as atrophic ARMD. In the recent article by Allikmets,15 most subjects with the ABCA4 variants had atrophic forms of ARMD. Yet other investigators have found no particular association of the G1961E allele with either exudative or atrophic ARMD.8 In the study by Allikmets, the classification of the patients was based on photographs in accordance with the International Classification System,16 which is designed to grade types and severity of ARMD-related disease but does not distinguish ARMD from other conditions. It is possible that a small subset of the “advanced ARMD” patients in that study actually had a typical or late-onset form of autosomal recessive Stargardt disease,17 since the natural history and age of onset of the condition prior to ascertainment for the study were not considered in the eligibility criteria. Even a relatively small number of these individuals could create a statistically significant association. As noted by Guymer and coworkers, most patients with Stargardt disease are compound heterozygotes, and one would not expect to detect many (if any) individuals who would be homozygous for the 2 alleles that were evaluated.15 Without exhaustive screening of the ABCA4 gene in the affected individuals harboring 1 copy of either G1961E or D2177N, one cannot exclude the possibility that their disease is actually an autosomal recessive form of Stargardt disease. Even this screening may be insufficient, given that we know that there are a relatively high number of patients with Stargardt disease for whom only a single ABCA4 mutation has been identified, despite complete sequence-based screening.8,18

Thus, the dilemmas and the controversy remain. Can one establish a universal set of criteria to define a causative gene for a complex disorder and, at the same time, can those criteria be realistically achieved? What does it mean to consider a gene as causing a complex disorder, rather than as a modifier or susceptibility gene that acts in concert with other specific genes and alleles? What level of penetrance do we demand before we can consider a variation in a gene to be responsible for a disease? In monogenic disorders, we typically consider penetrance to be 90% or higher. But what happens in a genetic model in which 50% or more of the people with the “disease-causing” allele may not develop the disease, perhaps because of the absence of an environmental trigger or because of modifier genes? What does causation really mean, or should we focus on the less satisfying and more vague concept of statistical risk assessments? One merely has to look at the digenic inheritance cases of retinitis pigmentosa19 or the genetics of multiple sclerosis or diabetes mellitus20 to get an initial glimpse of the confusing picture that can emerge.

What evidence should guide us toward implying the role of a particular gene with respect to a complex disorder? The article by Guymer et al.21 must give pause to our rush to champion the ABCA4 gene as the “first gene” responsible for ARMD. The hypothesis initially posed by Allikmets et al.15 must still be given its valid place in ARMD genetics, but its proof remains uncertain. In the absence of a validated animal model or a functional assay that can discriminate how specific variants of the ABCA4-related protein affect cellular function, we must continue to rely on association studies (both in families and populations) to identify candidate genes and to test our hypotheses. As described in this article, proof will be far more challenging and elusive than many of us initially suspected.

Michael B. Gorin, MD
Pittsburgh, Pa

REFERENCES