Common and Rare Variants in Alcohol Dependence

Howard J. Edenberg

Many lines of evidence converge to show that genetics makes a substantial contribution to the risk for alcohol dependence, explaining about 60% of the variance. However, alcohol dependence is not a Mendelian trait with a simple pattern of inheritance, nor is it deterministic. It is a complex trait like most psychiatric diseases and other common diseases, with both genetic and environmental factors affecting risk. It is difficult to identify genes that affect the risk for complex diseases, but the huge impact of these diseases makes the attempt important.

Early studies of alcohol dependence were limited to examining the role of small numbers of genetic variants in candidate genes, genes with a hypothesized role in alcoholism. Some were successful, particularly the identification of coding variations in *ADH1B* and *ALDH2*, genes that encode key enzymes in the metabolism of alcohol, with protective variants reducing risk by 2- to 8-fold (1). Although other genes have also been implicated, no other variants with comparable effects are known. Some candidate genes have remained controversial.

A strategy of linkage analysis followed up by genotyping of candidate genes within linked regions has also had success. A handful of genes have been confirmed to have variants affecting risk, including *GABRA2* and *ADH4* (2). Effect sizes are, however, small.

Technological progress in genotyping single-nucleotide polymorphisms (SNPs) triggered genome-wide association studies (GWAS), which have the potential to discover genes not previously thought to be involved. After an initial striking success for age-related macular degeneration (3), a wave of studies on many diseases was carried out with disappointing initial results: very few genes passed the accepted level of genome-wide significance, set at a high level of stringency because of the massive multiple testing these studies entail. Frequently, different studies on the same disease did not find the same results. Although this caused some to write off GWAS as a failure, a different take-home message began to emerge recently, as groups combined their data in metaanalyses. As the number of cases and controls grew dramatically, more and more genes have been identified in many diseases (4). It is clear that real effect sizes are much smaller than initially thought (or than calculated from the initial study, the so-called winners curse), so very large studies are needed to reliably detect them.

The results of several initial GWAS on alcohol dependence and related traits, including one in this issue (5), fit this pattern. The studies differ in populations and ascertainment. A study of German male alcoholics with early onset of the disease, ascertained from hospitals, did not yield any SNPs at genome-wide significance, but follow-up of the top SNPs and selected candidate genes showed nominal significance, and a combined analysis reported that two SNPs on chromosome 2q35 (in strong linkage disequilibrium) reached genome-wide significance (6). Edenberg *et al.* (7) carried out a GWAS on subjects from the Collaborative Studies on the Genetics of Alcoholism (COGA) study, with no SNPs reaching genome-wide significance. Combining gene expression data with the genetic data and a follow-up analysis in families, a region on chromosome 11 stood out, as did several individual genes. One of the top SNPs was in *BBX*, an HMG-BOX transcription factor whose expression is affected by ethanol. Six of the top SNPs from the German study (6) were at least nominally significant in the COGA study, with the same risk allele (7).

Bierut *et al.* (8) carried out a larger GWAS of alcoholic subjects and controls taken from three different studies, ascertained for alcohol dependence (a large subset of the sample in the other report [7]), cocaine dependence, and nicotine dependence. Again, no SNP reached genome-wide significance, and there was little overlap in top SNPs with either of the earlier GWAS.

A recent, large GWAS focused on a quantitative trait, alcohol consumption (grams per day per kilogram body weight), rather than alcohol dependence (9). It was assembled from many different studies, and the level of drinking for most subjects was modest. A SNP in *AUTS2* reached genome-wide significance, and functional studies provided support for its involvement in alcohol-related phenotypes.

In this issue, Heath *et al.* (5) report results from a GWAS on Australian subjects recruited from several samples. The cases in this study, recruited from the community, were on average much less severe than the cases recruited from treatment facilities in the above-mentioned studies of alcohol dependence (6–8). This difference is likely to affect the genes identified. This study also differed from the previous ones by using family-based analyses rather than case-control.

Another major difference is that in addition to analyzing alcohol dependence, several quantitative traits were constructed: factor scores for heaviness of drinking and for alcohol use disorders. Despite the careful analysis of both categorical and quantitative traits, no SNPs reached genome-wide significance, although some were suggestive. It is interesting that a set of SNPs in *BBX*, associated with dependence in the COGA study (7), were associated with several consumption measures.

Key findings from the article by Heath *et al.* (5) are that for alcohol dependence and the quantitative drinking traits they analyzed, the effect size of individual variants was estimated to be very small, in the range of 0.25% and below, and there are likely to be hundreds of genes of small effect. This suggests that much larger studies and metaanalyses will be needed to reliably identify genes in which there are variants affecting the risk for alcohol dependence. To date, the number of subjects with data on alcohol dependence that have been studied is low compared with other diseases. Several GWAS are in progress, which will help once they can be combined in meta-analyses, but far more subjects assessed for alcohol dependence will be needed. Although it will be easier to collect studies with data on alcohol consumption, most were not targeted at problem drinking and will address only one aspect of the problem and not the genetics of dependence.

There is currently a new focus on rare variants. GWAS target common alleles. The SNPs chosen for GWAS arrays generally have high allele frequencies, and most analyses are restricted to minor allele frequencies greater than 1%. This is both a practical matter, because common SNPs give much more power, and based on the common disease–common variant hypothesis that much of the genetic risk for common diseases is due to loci at which a single

From the Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana.

Address correspondence to Howard J. Edenberg, Ph.D., Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, MS4063, Indianapolis, Indiana 46202-5122; E-mail: edenberg@iupui.edu.

Received July 15, 2011; accepted July 15, 2011.

0006-3223/$36.00
doi:10.1016/j.biopsych.2011.07.007

common variant is the main contributor. If there is allelic heterogeneity, the power of GWAS declines dramatically.

An alternative hypothesis is that many rare variants, each of large effect but acting in only a small fraction of the cases, make a substantial contribution to the genetics of common disease (common disease–many rare variants). Rare variants are not covered by GWAS arrays, although Goldstein (10) has proposed that some of the signals from common SNPs actually represent synthetic associations due to multiple rare alleles in different degrees of linkage disequilibrium with the SNPs on the arrays.

Because rare variants differ in different subjects and most have probably not yet been described, several groups are turning to next-generation sequencing to identify them. They may be missense or nonsense mutations or alterations in splice sites or key regulatory sites that dramatically affect the structure and function of the protein encoded by the gene. There are so many genetic differences among individuals that the identification and confirmation of potentially causative variants is quite challenging. Family studies, which dropped out of favor in the initial enthusiasm for GWAS, provide an excellent resource for this because one can test whether a proposed causal variant segregates within the family in a manner consistent with the disease.

Obviously the common disease–common variant hypothesis and the common disease–many rare variants hypothesis are not mutually exclusive. In fact, biology strongly suggests that if variation in a gene has an impact on a biological process or disease, there will be a spectrum of variations with a spectrum of effects, including common variants of small effect and rare variants of large effect. That, and the potential contribution of synthetic associations, both argue for a strategy of deep sequencing genes for which there is evidence of common alleles affecting the disease to determine whether there are also rare variants of large effect in particular families. Another potentially powerful strategy to identify important rare variants is to carry out whole-genome sequencing in key members of large families in which there is evidence for linkage, and focus analyses on functional variants found within the linkage peaks.

There is an area in which I am more optimistic than Heath *et al.* (5). They argue that because effect sizes for genes found in GWAS are so small, they may not be of major benefit to drug discovery. However, they can give crucial leads in two ways. First, the genes identified in GWAS can be targeted for studies of rare variants. Second, the effect size of a common variant does not determine the impact of targeting that gene with a drug. Even if the variant discovered has a small effect, the effects of a drug targeting the gene (or another step in the pathway in which it functions) could nevertheless be large. Thus, more studies, both GWAS and sequencing, are critical if we are to progress in our understanding of the disease and our ability to better treat patients.

Related work in the author's laboratory is supported by National Institute on Alcohol Abuse and Alcoholism Grants AA008401 and AA006460. The funding agency had no role in the preparation of this manuscript. The author reports no biomedical financial interests or potential conflicts of interest.

1. Edenberg HJ (2007): The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res Health* 30:5–13.
2. Edenberg HJ, Foroud T (2006): The genetics of alcoholism: identifying specific genes through family studies. *Addict Biol* 11:386–396.
3. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, *et al.* (2005): Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–389.
4. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A Catalog of published genome-wide association studies. Available at: http://www.genome.gov/gwastudies.
5. Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA, *et al.* (2011): A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* 70:513–518.
6. Treutlein J, Cichon S, Ridinger M, Wodarz N, Soyka M, Zill P, *et al.* (2009): Genome-wide association study of alcohol dependence. *Arch Gen Psychiatry* 66:773–784.
7. Edenberg HJ, Koller DL, Xuei X, Wetherill L, McClintick JN, Almasy L, *et al.* (2010): Genome-wide association study of alcohol dependence implicates a region on chromosome 11. *Alcohol Clin Exp Res* 34:840–852.
8. Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, *et al.* (2010): A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci USA* 107:5082–5087.
9. Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D, *et al.* (2011): Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proc Natl Acad Sci USA* 108:7119–7124.
10. Goldstein DB (2009): Common genetic variation and human traits. *N Engl J Med* 360:1696–1698.