cmgh
CELLULAR AND MOLECULAR GASTROENTEROLOGY AND HEPATOLOGY

EDITORIAL

Digesting GWAS

Once upon a time, medical genetics was a relatively straightforward business: one gene—one protein—one disease was sufficient to explain most inherited disease and was neatly in keeping with Gregor Mendel's discovery of alleles with his famous peas. The concept remains a good starting point but the biology of life is somewhat more complex, and increasingly we have become impressed and perplexed, in equal measure, at the striking variability observed in many inherited conditions, particularly those hitherto believed to be determined by a heterozygous single-gene mutation or DNA variant, and therefore following autosomal-dominant inheritance. A good example is Alagille syndrome, and in this issue of *Cellular and Molecular Gastroenterology and Hepatology*, Tsai et al[1] proposed a modifier genetic locus to help explain the great variation in severity of liver disease in this condition. The authors reached this conclusion by using a technique that is barely more than 10 years old, namely a *genome-wide association study* (GWAS), which is part of the arsenal of techniques that has heralded the era of genomic medicine.

A GWAS, rather similar to DNA fingerprinting, exploits the immense variation in DNA sequence between individuals, these differences consisting of single-nucleotide polymorphisms (SNPs), deletions, insertions, and copy number variations, the majority of which are regarded as essentially benign. Although a genetic linkage study will set out to identify the rare single-gene mutations or variants that clearly cause a particular disease or phenotype, an association study is potentially useful at finding common genetic variants with a relatively weak effect in more complex disorders. A typical study will adopt a case–control model, comparing 2 cohorts of subjects, 1 with the phenotype or disorder of interest and the other functioning as normal controls. It goes without saying that the criteria for selecting the cohort with the phenotype should be clinically rigorous whereas the control group should be as pure as possible (ie, free of contamination from subjects who might harbor the phenotype). Generally speaking, the larger the number of subjects in each group, the better the chances of achieving significant results. The genomes of all subjects then are analyzed for the presence of at least a million common SNPs, after which the investigation moves to the statistical phase: are there certain SNPs or alleles that are seen more often in the disease group compared with controls, and, if so, by how much (ie, what is the odds ratio?).

This is all well and good, but there are pitfalls and shortcomings aplenty. Apart from the crucial aspect of carefully selecting the study cohort by clinical phenotype, the population background should be taken into account lest there are common origins with a range of unique but misleading polymorphisms that give rise to false associations. The normal variation data accumulating through many large-scale, biobank-style initiatives and the Human Variome Project will be vital to the interpretation of results in many areas of clinical genomics—an issue referred to as *population stratification*. Ideally, studies should be reproducible in more than one population. Then it must be appreciated that the results of a GWAS identify a genetic locus or high-risk SNPs *associated* with the disease in question—which is unlikely to be the actual *causal* gene or DNA variation. Therefore, beyond the odds ratio, it is of course pertinent to ask what difference a significant association will make to clinical management. Unless further fine mapping and sequencing work is performed to identify the specific gene or DNA variation that causes or modifies the phenotype, the GWAS statistical results from a study cohort cannot be meaningfully applied to an individual in a clinical setting, all of which is highlighted by the study by Tsai et al.[1]

Perhaps the most fundamental issue is the starting assumption that a common medical condition, or the variability of a relatively rare genetic disease, has a significant genetic basis that will be uncovered by GWAS. These studies are expensive and can be criticized on the grounds that few of them have led to any tangible clinical benefit. However, some have, particularly in relation to identifying the variation in response to treatment for various diseases, eg, hepatitis C, in which SNPs close to the interleukin 28B gene determined this variation,[2] and the risk of pancreatitis related to treatment of inflammatory bowel disease by thiopurine immunosuppressants.[3]

GWA studies will continue to have a place in researching genetic variation and susceptibility to disease, although increasingly will be combined with the next-generation sequencing strategies of whole-exome and whole-genome sequencing, the costs of which continue to fall. The hope and expectation is that these technologies will spearhead advances that will deliver the much-anticipated era of personalized medicine. The article by Tsai et al,[1] describing THBS2 as a candidate modifier of liver disease severity in Alagille syndrome, represents the beginning of a process that in due course may lead to a change in the way patients with Alagille syndrome are evaluated and screened.

PETER D. TURNPENNY, FRCP, FRCPCH, FRCPath
University of Exeter Medical School
Clinical Genetics
Royal Devon and Exeter Hospital
Exeter, United Kingdom

References

1. Tsai EA, Gilbert MA, Grochowski CM, et al. THBS2 is a candidate modifier of liver disease severity in Alagille syndrome. Cell Mol Gastroenterol Hepatol 2016;2:663–675.
2. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461:399–401.

3. Heap GA, Weedon MN, Bewshea CM, et al. HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants. Nat Genet 2014;46:1131–1134.

Correspondence
Address correspondence to: Peter D. Turnpenny, MD, University of Exeter Medical School, Clinical Genetics, Royal Devon and Exeter Hospital, Gladstone Road, Exeter EX1 2ED, United Kingdom. e-mail: peter.turnpenny@nhs.net.

Conflicts of interest
The author discloses no conflicts.

Most current article

© 2016 The Author. Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
2352-345X
http://dx.doi.org/10.1016/j.jcmgh.2016.07.003