For those who have witnessed the evolution of genetic association studies of plasma lipids since 1983, the field has had its roller-coaster moments. The early high points followed from the application of “restriction fragment length polymorphism” analysis to yield candidate genotypes that were statistically evaluated for association with plasma lipoproteins. After 2 decades, though, the candidate gene approach had produced only negligible replicable associations, notably with genotypes for the canonical APOE isoforms and some others. Because of this inconsistency, enthusiasm for candidate gene association studies had essentially been exhausted by 2005. However, in late 2007, association studies experienced a renaissance and have generated important new findings for lipoprotein metabolism.

The reasons underlying this resurrection have been widely discussed and include both new technology—namely, microarrays to genotype hundreds of thousands of single-nucleotide polymorphisms (SNPs)—plus very large sample sizes, permitting metaanalysis of individual-level or summary data from tens of thousands of individuals. Over the last 3 years, genome-wide association studies (GWASs) have identified robust, replicable statistical associations for numerous complex diseases and traits, including plasma lipids. The basic GWAS design and workflow are shown in the Figure. The study by Waterworth et al in the current issue of Arteriosclerosis, Thrombosis, and Vascular Biology illustrates both strengths and limitations of the GWAS approach.

The authors report metaanalyses of genotype associations with plasma lipids from >17,000 white European subjects, by combining GWAS data from 8 studies, followed by replications in >37,000 individuals from 8 additional studies and in a smaller sample of South Asian subjects. Significant SNP genotypes were then examined for their association with coronary artery disease risk in >9,600 cases and >38,000 controls.

The findings confirm past associations of low-density lipoprotein–cholesterol (LDL-C), high-density lipoprotein–cholesterol, and triglyceride concentrations with 9, 10, and 9 loci, respectively. Several loci showed pleiotropic associations with ≥2 traits. Similar, albeit fewer significant associations were found in the smaller South Asian sample. The authors emphasized 4 new replicated genetic loci: MYLIP/GMPR and PPP1R3B for LDL-C, SLC39A8 for high-density lipoprotein–cholesterol, and AFF for triglyceride. They showed that ≈1/4 of the lipid-associated loci were also associated with coronary artery disease risk.

However, despite its strong design and its interesting findings, this study is probably among the last of a dying breed. The landscape for reporting GWASs of plasma lipids changed forever in August 2010. Going forward, all future reports of genetic determinants of plasma lipids should be considered in light of the major achievement of organization and collaboration that characterized the Global Lipids Genetics Consortium (GLGC) metaanalysis.

In GLGC, >200 collaborators pooled results in which they screened the genomes of >100,000 white Europeans with ≈2.6 million common SNP genotypes and tested them for associations with plasma lipids. They reported 95 significantly associated loci, of which 59 were completely novel. The new loci contained genes that were both known to have roles in lipoprotein metabolism from nongenetic research but also contained previously unknown genes.

The same 95 loci contributed to variation in lipid traits from both population-based epidemiological samples and to severe dyslipidemia phenotypes. Studies in 3 non-European ethnic groups showed similar associated loci. In a case-control analysis, almost all LDL-C-associated SNPs and a smaller proportion of triglyceride and high-density lipoprotein–cholesterol–associated SNPs were associated with coronary artery disease risk. Finally, functional studies validated 3 novel genes—GALNT2, PPP1R3B, and TTC39B—as playing mechanistic roles in determining plasma lipid concentrations.

The GLGC report had all the desirable attributes of a comprehensive genetic association study, including huge numbers of study subjects drawn from multiple diverse samples, sustainable replications, and demonstration of a mechanistic impact of some genes and polymorphisms. The main findings from Waterworth et al were also seen in GLGC, including all 4 new replicated loci, although the triglyceride locus marked by AFF by Waterworth et al was designated KLHL8 in GLGC. The overlap of findings between the 2 studies is unsurprising: most samples (and authors) from Waterworth et al were subsumed into GLGC. Another interesting wrinkle found by Waterworth et al was attribution of an association signal with LDL-C to CELSR2. In contrast, GLGC and other studies have implicated the neighboring gene SORT1 (encoding sortilin 1) as the source of the association signal at this locus; functional studies in a companion article to the GLGC report provide mechanistic...
data further supporting a causal role for \textit{SORT1} as the basis for the association signal. Interestingly, the locus also harbors another gene, \textit{PSRC1}; perhaps all 3 genes somehow contribute to the LDL-C association. This will be an interesting story to monitor.

With the GLGC, GWASs of lipids in whites may have reached the limit of what can be practically expected. The enormous statistical power in GLGC permitted sensitive detection of some loci with exquisitely small genetic effects. Although these loci may prove to be biologically relevant, the
clinical or diagnostic relevance of any specific variant may be harder to demonstrate. Cumulatively, ≈30 SNP genotypes explain ≈10% of the total variation (and ≈25% of variation attributable to genetic factors) for each lipid trait.\(^4\)

An important lingering question regards the source or sources of the remaining unattributed or “missing” variability. One possibility seems to be low frequency genetic variants with large biochemical effects that cannot be detected with DNA microarrays. Such variants would require different types of arrays\(^7\) or DNA sequencing to be detected. In hypertriglyceridemia, rare heterozygous variants have already been shown to explain some of this missing genetic component.\(^8\)

Thus, as quickly as they illuminated the landscape, GWASs of plasma lipids may soon begin to fade from the scene. Future studies will include the following features: (1) assessing additional geographical ancestries; (2) evaluating rare genetic variants; (3) screening larger cohorts with clinical dyslipidemia; (4) using high-throughput functional characterization of gene products; (5) evaluating nonlinear and nonadditive genetic models; and (6) assessing gene-gene and gene-environment interactions. However, much of the groundwork has now been laid by reports like those from Waterworth et al and the GLGC.\(^3\) It will be exciting to follow the progress of GWAS data as they begin to have an impact on biological and medical research.

Acknowledgments

Christopher Johansen provided helpful comments and assisted with the figure preparation. The author is supported by operating grants from the Heart and Stroke Foundation of Ontario (NA 6018 and NA 6059), the Canadian Institutes for Health Research (MOP 13430 and 79533), and Genome Canada through the Ontario Genomics Institute.

References

Genome-Wide Association Studies of Plasma Lipids: Have We Reached the Limit?

Robert A. Hegele

Arterioscler Thromb Vasc Biol. 2010;30:2084-2086
doi: 10.1161/ATVBAHA.110.214643

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://atvb.ahajournals.org/content/30/11/2084

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/