Does a Gene Polymorphism Predisposing to an Intermediate Phenotype Predict the Risk of Disease?  
— A Lesson From CETP, High-Density Lipoprotein-Cholesterol and Coronary Artery Disease ——

Akinori Kimura, MD*,**

Prevention of disease is the ultimate goal of medicine and to realize it, many step-by-step approaches have been taken. One approach is the identification of clinical parameters closely associated with the disease, because such parameters or intermediate phenotypes might be useful for predicting the risk of disease and, more importantly, might be interfered with to prevent the disease. For example, hyperlipidemia, or more specifically hypercholesterolemia, is a well-known clinical parameter closely associated with atherosclerosis, including coronary atherosclerosis, and hence associated with coronary artery disease (CAD), including myocardial infarction (MI). In other words, hypercholesterolemia and atherosclerosis can be intermediate phenotypes of CAD (Figure). Although there are several other risk factors for CAD (e.g., smoking, hypertension, obesity and diabetes mellitus), which should be handled in appropriate ways to achieve the prevention of CAD, hyperlipidemia is a target for treatment and this is the rationale for cholesterol-lowering medication being used as a tool for preventing CAD in the clinics. It is also important to decipher the molecular mechanisms of regulation of the serum cholesterol level and its involvement in the pathogenesis of atherosclerosis in order to develop novel diagnostic or pharmacological tools in the management of CAD.

Another stepwise approach taken in the past decade is the identification of genetic risk factors predisposing to the development of disease, because genetic factors are more or less involved in the etiology and/or pathogenesis of disease. Familial aggregation of CAD has long been known and it is reported that the contribution of genetic factors in CAD can be represented as an odds ratio of approximately 2 to 4. In addition, a twin study demonstrated that the risk of CAD-related death in male siblings of the affected proband aged 36–65 years was 15.0 and 2.6, respectively. These data apparently show the considerable contribution of genetic factors to CAD, in addition to environmental or lifestyle factors, because the difference between monozygotic and dizygotic twins should mainly represent genetic differences, not differences in environmental factors.

As a result of the human genome project, diversity in the human genome has now been clarified in detail, particularly the many single nucleotide polymorphisms (SNP) and copy number variations (CNV), which are useful genetic markers to be investigated in association with the diseases or intermediate phenotypes. Indeed, an increasing number of papers have reported the association of specific variations in the human genome with the intermediate phenotypes of CAD, dyslipidemia or atherosclerosis, including some reports published recently in the Circulation Journal. Most of the initial association studies have focused on the polymorphisms of candidate genes, because a priori knowledge is useful in interpretation of data that would be obtained from the association studies. On the other hand, genome-wide association (GWA) studies of the SNPs and CNVs with the disease have recently been conducted, because GWA studies can be expected to identify the disease-related human genome diversity without prior knowledge about the disease.

The opinions expressed in this article are not necessarily those of the editors or of the Japanese Circulation Society.

*Department of Molecular Pathogenesis, Medical Research Institute, **Laboratory of Genome Diversity, School of Biomedical Science, Tokyo Medical and Dental University, Tokyo, Japan

Mailing address: Akinori Kimura, MD, Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

All rights are reserved to the Japanese Circulation Society. For permissions, please e-mail: cj@j-circ.or.jp

Figure. Contribution of life-style factor and genetic factors in determining intermediate phenotypes and disease. Different genetic factors may participate in each step of developing coronary artery disease. HDLC, high-density lipoprotein-cholesterol; LDLC, low-density lipoprotein-cholesterol.
In this issue, Hiura et al reported the results of a middle-scale GWA study for high-density lipoprotein-cholesterol (HDL-C) levels in Japanese. They investigated 900 Japanese subjects for 368,274 SNPs and found 43 SNPs showing P<1.0E-4 (arbitrarily determined significance level), including a SNP (rs3764261) in the 5' upstream of CETP with a P-value less than 1.0E-6. Although the P value for rs3764261 did not reach statistical significance after correction of multiple testing, by increasing the number of subjects they showed that the copy number of the rs3764261 minor allele was significantly associated with an increase in HDL-C by 6.2 mg/dl! This is not surprising because there have been several reports of the association of SNPs in CETP with HDL-C level in Caucasian populations. In addition, 2 quite large-scale GWA studies have recently been reported. One study investigated 17,797–22,562 persons in European populations and the other study conducted a study of 19,840 individuals for screening and 20,632 individuals for replication. The former study identified 9 gene loci that were significantly (P<5.0E-8) correlated with HDL-C level, while the latter identified 14 gene loci showing a significant association (P<5.0E-8). Although the CETP locus was ranked at the top significance in both studies (9.4E-94 and 4.0E-75, respectively), it should be noted that only 5 gene loci (ABCA1, CETP, LIPC, LIPG and LPL) were identified as significantly associated with HDL-C level in both studies. The other loci showing significant association in only 1 study were considered to play a role in HDL-C level, but their contribution was not large enough to be replicated by the other study. In turn, these data suggest that much larger samples are required to capture the loci with a small contribution. In this regard, while Hiura et al also reported a suggestive linkage of HDL-C level with the FLJ45139-ETS2 loci on chromosome 21q22, it was not identified by the large-scale GWA studies in Caucasian populations, suggesting that the loci proposed by Hiura et al should be confirmed by other studies before reaching a definite conclusion about its contribution to HDL-C.

On the other hand, the GWA approach has been used to identify disease-associated loci. Large-scale GWA studies in European populations identified several loci with a significant association in each study, but only 1 locus on chromosome 9p21 was independently replicated in those studies. The association of CAD with the 9p21 locus has been replicated in Japanese and Korean populations, suggesting that the loci proposed by Hiura et al should be confirmed by other studies before reaching a definite conclusion about its contribution to HDL-C.

In this issue, Hiura et al investigated their data in association with HDL-C level, while the latter identified 14 gene loci showing a significant association (P<5.0E-8). Although the CETP locus was ranked at the top significance in both studies (9.4E-94 and 4.0E-75, respectively), it should be noted that only 5 gene loci (ABCA1, CETP, LIPC, LIPG and LPL) were identified as significantly associated with HDL-C level in both studies. The other loci showing significant association in only 1 study were considered to play a role in HDL-C level, but their contribution was not large enough to be replicated by the other study. In turn, these data suggest that much larger samples are required to capture the loci with a small contribution. In this regard, while Hiura et al also reported a suggestive linkage of HDL-C level with the FLJ45139-ETS2 loci on chromosome 21q22, it was not identified by the large-scale GWA studies in Caucasian populations, suggesting that the loci proposed by Hiura et al should be confirmed by other studies before reaching a definite conclusion about its contribution to HDL-C.

In this issue, Hiura et al reported the results of a middle-scale GWA study for high-density lipoprotein-cholesterol (HDL-C) levels in Japanese. They investigated 900 Japanese subjects for 368,274 SNPs and found 43 SNPs showing P<1.0E-4 (arbitrarily determined significance level), including a SNP (rs3764261) in the 5' upstream of CETP with a P-value less than 1.0E-6. Although the P value for rs3764261 did not reach statistical significance after correction of multiple testing, by increasing the number of subjects they showed that the copy number of the rs3764261 minor allele was significantly associated with an increase in HDL-C by 6.2 mg/dl! This is not surprising because there have been several reports of the association of SNPs in CETP with HDL-C level in Caucasian populations. In addition, 2 quite large-scale GWA studies have recently been reported. One study investigated 17,797–22,562 persons in European populations and the other study conducted a study of 19,840 individuals for screening and 20,632 individuals for replication. The former study identified 9 gene loci that were significantly (P<5.0E-8) correlated with HDL-C level, while the latter identified 14 gene loci showing a significant association (P<5.0E-8). Although the CETP locus was ranked at the top significance in both studies (9.4E-94 and 4.0E-75, respectively), it should be noted that only 5 gene loci (ABCA1, CETP, LIPC, LIPG and LPL) were identified as significantly associated with HDL-C level in both studies. The other loci showing significant association in only 1 study were considered to play a role in HDL-C level, but their contribution was not large enough to be replicated by the other study. In turn, these data suggest that much larger samples are required to capture the loci with a small contribution. In this regard, while Hiura et al also reported a suggestive linkage of HDL-C level with the FLJ45139-ETS2 loci on chromosome 21q22, it was not identified by the large-scale GWA studies in Caucasian populations, suggesting that the loci proposed by Hiura et al should be confirmed by other studies before reaching a definite conclusion about its contribution to HDL-C.

On the other hand, the GWA approach has been used to identify disease-associated loci. Large-scale GWA studies in European populations identified several loci with a significant association in each study, but only 1 locus on chromosome 9p21 was independently replicated in those studies. The association of CAD with the 9p21 locus has been replicated in Japanese and Korean populations, suggesting that the loci proposed by Hiura et al should be confirmed by other studies before reaching a definite conclusion about its contribution to HDL-C.