A Strategy for Genomic Research on Common Cardiovascular Diseases Aiming at the Realization of Precision Medicine

Personal Insights and Perspectives

Hiroyuki Morita, Issei Komuro

Despite the development of genomic research on common cardiovascular diseases, genetic variants and/or loci shown to be disease associated have been seldom applied to risk prediction in a clinical setting, although they could provide us with novel mechanistic insights into the pathophysiology and potential therapeutic targets. Here, we discuss the significance and limitations of the up-to-date genomic studies on common diseases and propose what kind of genomic research should be prioritized in view of the realization of precision medicine.

Many disease-associated loci have been identified using a genome-wide association study (GWAS) in a large-scale case–control design (more than several thousand cases and controls). A GWAS is an analytical method for the identification of disease-associated tag (lead) SNPs, which are indicators of disease-associated loci (linkage disequilibrium blocks) with a genome-wide significance ($P<5\times10^{-8}$). Consequently, further analysis is required to identify the specific variants functioning as the most associated and/or risk variants[1] within each disease-associated locus. Consistent with the genetic cause of monogenic diseases, novel biology on the risk genes and their regulatory pathways can be learned through detailed genetic and functional analysis after GWAS. Also, GWAS data on disease-associated pathways might be used in choosing drugs for treatment.[2] We have experienced several successful stories. Single-nucleotide polymorphisms (SNPs) in *PCSK9*, *HMGCR*, and *NPC1L*1 were shown to be associated with hypercholesterolemia in GWAS on blood lipids.[3] Although their effects shown in that GWAS were relatively weak (+2–3 mg/dL change in plasma cholesterol values), potent compounds targeting these gene-encoded proteins have been established as lipid-lowering agents (anti-PCSK9 antibody, statin, and ezetimibe, respectively). In this way, GWAS data can provide the scientific community with information about novel biology, mechanical pathways, and potential pharmaceutical drugs.

Ideal Genomic Research Project to Overcome the Limitation of Current GWASs

Because current GWASs are usually performed with a case–control design, the results are difficult to apply directly to the prediction tool for prospective onset in the general population or future recurrence in patients. More importantly, most of the current GWASs were not originally designed to obtain genetic information useful for the stratification of individuals according to the severity, prognosis, and responsiveness to therapies but rather to clarify the genetic variants related to the presence of disease. Therefore, results from the current GWASs by themselves are only minimally applicable for individual predictions of clinical outcomes.

What should we do from an idealistic viewpoint? At first, we should compile clinical questions that genetic testing can effectively answer and select the most attractive questions from among them. In particular, we should prioritize the clinical questions that cannot be answered only using the clinical or laboratory examinations available. Specifying the questions to be answered, we can set the subpopulations according to the clinical subphenotype of interest in the study population. The analysis on the variants related to the subphenotypes will provide us with the more detailed clinical information useful for the classification of patients and the individual prediction of their clinical outcomes. Within the patients with coronary artery disease (CAD), for example, we can set the subpopulations as follows: acute coronary syndrome (versus stable angina), early-onset, multiple atherosclerotic lesions, ischemia-induced cardiac remodeling or heart failure, and the patients susceptible to cardiometabolic risks (as described below). Such a variant associated with subphenotype within a disease phenotype could function as a second-hit, independently of the variant associated with the disease onset per se. It remains unknown whether its genetic impact might be smaller or larger as compared with that of disease-onset–associated variants.

Taken together, we should conduct large-scale prospective studies originally designed to compare the distributions of genetic variants among the subpopulations stratified according to subphenotype even in patients with similar clinical presentations because this should contribute to the realization of precision medicine. A large number of genomic samples should be obtained from a cohort study and disease registry, in which detailed clinical information is being prospectively collected. Additionally, to

From the Department of Cardiovascular Medicine, University of Tokyo, Japan.

Correspondence to Issei Komuro, MD, PhD, Department of Cardiovascular Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655, Japan. E-mail komuro-tky@umin.ac.jp

(*Circ Res*. 2016;119:900-903.
DOI: 10.1161/CIRCRESAHA.116.309802.)

***Circulation Research* is available at http://circres.ahajournals.org**
DOI: 10.1161/CIRCRESAHA.116.309802

Nonstandard Abbreviations and Acronyms	
CAD	coronary artery disease
GWAS	genome-wide association study
SNP	single-nucleotide polymorphism

acquire the most useful findings, we have to bring together the essence of next-generation sequencing and multiomics from the genome, epigenome, transcriptome, proteome, and metabolome to the phenome. The huge volume of genomic data, clinical data, and the enormous complexity of their interplay should be effectively analyzed, integrated, and submitted into the public databases. This should be a mainstay for upcoming genomic research toward the realization of precision medicine (Figure).

Identification of Risk Variants That Could Indicate Effective Clinical Action

As the first step toward the realization of precision medicine, for the present, we should prioritize the identification of risk variants, which could prompt clinical practitioners to take effective clinical action to the carrier of that particular variant (eg, preemptive intensive risk control, avoidance of medication).

With genomic analysis based on the hypothesis that several cardiometabolic traits related to cardiovascular risk factors might partly arise from a shared underlying genetic basis with CAD, 67 novel loci associated with CAD were identified.[4] Genetic variants ascertained as having an effect on cardiometabolic traits were shown to have correlated effects on risk of CAD. These genetic correlations show that these cardiometabolic traits (eg, low-density lipoprotein cholesterol and triglycerides) function as an established risk for the onset of CAD, as was shown in the previous Mendelian randomization studies.[5] From a different viewpoint, this study[4] successfully showed that the susceptibility to a cardiometabolic abnormality and resultant onset of CAD could be ascribed to genetic predispositions. In other words, genetic variants do not necessarily function as a risk for CAD independently of the conventional cardiometabolic risks. Further analysis in the subpopulations (eg, dyslipidemic patients with CAD versus those without CAD) is warranted to identify the genetic variants related to the susceptibility to CAD in the individuals with cardiometabolic abnormality. Such genetic variants will enable us to choose individuals susceptible to CAD, who should receive the preemptive intensive management of cardiometabolic abnormality in order not to have CAD.

Pharmacogenomics is also the clear choice for enhancing the near-term impact of precision medicine. The magnitude of the pharmacogenomic effects is typically larger than that of the individual variant effects on the disease. In the cardiovascular field, a vast amount of pharmacogenomics research on warfarin, clopidogrel, and statins has been conducted, and the genetic impact on drug efficacy has been clarified

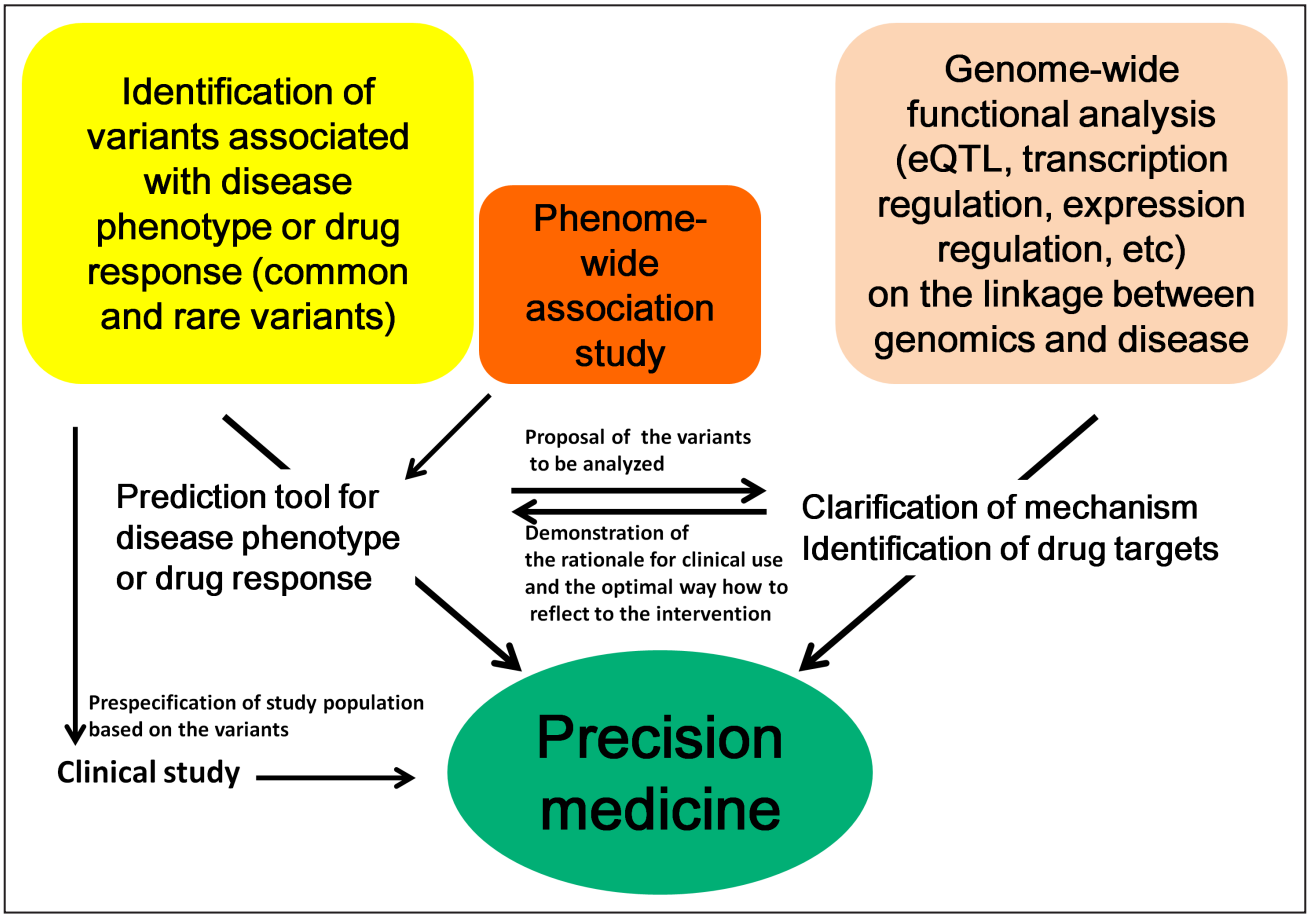

Figure. Schematic representation of genomic research on common diseases toward the realization of precision medicine.

but remains controversial. Rather, we had better prioritize the identification of genetic variants related to the susceptibility to drug-induced side effects. For example, a common variant in *SCLO1B1* was shown to be strongly associated with statin-induced myopathy.[6] Recently, a variant in the *RARG* (retinoic acid receptor gamma) gene was reported to be associated with the susceptibility to anthracyclin-induced cardiotoxicity.[7] *Top2b*, encoding topoisomerase IIβ, is necessary for the development of anthracyclin-induced cardiotoxicity. *Top2b* expression is basically repressed by RARG. The variant-type RARG cannot repress *Top2b* expression as effectively as wild-type RARG, leading to increased susceptibility to anthracyclin-induced cardiotoxicity. Basically, genetic tests identify the patients who would derive a differential net benefit from the drug therapy of interest. Nevertheless, at present, most clinical studies on drug therapy do not include pharmacogenomic data. Adding pharmacogenomic data to the variables for prespecification in the clinical studies, responders and nonresponders can be well prespecified, and consequently, the quality of clinical studies will be improved in that more accurate comparisons could be performed.[8]

Special Considerations for the Clinical Usefulness of Common Variants

The issues above shall not apply to all variants associated with common phenotype. Because of the inherent nature of common variants exerting small effect sizes (with odds ratio of <1.5), each variant explains only a small fraction of the variance in the clinical phenotype. Actually, skepticism that common variants might not be useful for the prediction of clinical phenotype has been argued. To overcome this concern and build up the excellent algorithm for risk prediction, the following 3 issues are required. (1) A compendium of common variants has to be comprehensively clarified. All associated variants should be fully identified as soon as possible. (2) Rarer variants that have a greater functional consequence in an individual carrier than common ones should be also identified. Indeed, rare variants were shown to be associated with CAD.[9,10] (3) We have to elaborate a method for establishing a good algorithm for risk prediction. To date, a weighted genetic risk score has been constructed by summing the number of risk alleles for each SNP weighted by their estimated effect sizes in the discovery sample. As the relationship among SNPs can now be more profoundly understood using newly developing studies including phenome-wide association study[11] and mathematical modeling, hereafter, the method to choose and combine SNPs can be improved.

In a good contrast with the rarer variants or mutations causing monogenic disorder, common variants can explain a larger proportion of clinical variation at the population level because of their high frequency. Given that, genomic research focusing on the clinical application of genomic findings on common variants to risk prediction is essential for the realization of precision medicine.

Genome-Wide Functional Analysis Is Essential for the Clinical Use of Genomic Findings

Some CAD GWAS loci are associated with the known risk factors for CAD,[4] but the underlying pathophysiologies of the remaining ones are still unknown. Most variants with the lowest *P* values fall in noncoding intergenic regions, and basic research is required to clarify their functional consequences to the disease. Without knowledge of these functional consequences, the application of these disease-associated variants into precision medicine cannot be practically achieved. In other words, clarification of the functional mechanism linking genetic variants to diseases is a limiting step for their clinical use as a prediction tool.

Recently, genetic loci can be functionally annotated using genome-wide methods with reference to developing public databases, which helps prioritize loci with likely biological function. Regulatory annotations of common genetic variants using multiple ENCODE (Encyclopedia of DNA Elements) data sets (DNase hypersensitivity, histone ChIP-seq data, and transcription factor ChIP-seq data) can contribute to the identification of causal variants and the prediction of sequence-encoded regulation of gene expression. At GWAS-identified susceptible loci for type 2 diabetes mellitus, the variants driving association signals are enriched for overlap with the transcription factor FOXA2-binding sites assayed by ChIP-seq, and the *MTNR1B* variant located in the FOXA2-bound enhancer causes the higher expression of *MTNR1B*, leading to the susceptibility to type 2 diabetes mellitus.[12] In another report, a variant in the *FTO* region associated with obesity was found to disrupt ARID5B-mediated repression of the downstream target genes *IRX3* and *IRX5*, followed by activation of pro-obesity pathways. Repair of the ARID5B motif by CRISPR-Cas9 editing of this predicted causal variant in primary adipocytes from a patient restored *IRX3* and *IRX5* repression, exerting a significant effect on obesity phenotypes.[13] Methylation analysis could also show the regulatory mechanisms underlying gene expression. Increased methylation at the *HIF3A* locus, which was inversely correlated with *HIF3A* gene expression in adipose tissue, was shown to be associated with increased body mass index.[14]

CAD is a complex disease including multiple tissues/cells. The use of expression data from several tissues/cells involved in CAD could lead to an incorrect interpretation of the results from eQTL (expression quantitative trait loci) analysis. And, because of the inaccessibility of tissues/cells involved in CAD, the eQTL analysis on CAD is difficult to perform. Under such a situation, a recent study successfully demonstrated a link between an intronic SNP in *PHACTR1*, transcription factors myocyte enhancer factor-2 binding in vascular endothelial cells, *PHACTR1* expression levels in coronary arteries, and CAD risk.[15] To clarify the functional consequences of common variants, comprehensive analysis should be performed using biological samples from biorepositories (biobanks) that assemble, store, and manage collections of human specimens and related data.

Conclusions

In parallel with the unbiased, hypothesis-free, comprehensive identification of common phenotype-associated variants, we should establish algorithms for risk prediction and perform functional analysis to effectively use the achievements of genomic research on common cardiovascular diseases for the realization of precision medicine.

Disclosures

None.

References

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424. doi: 10.1038/gim.2015.30.
2. Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, Floratos A, Sham PC, Li MJ, Wang J, Cardon LR, Whittaker JC, Sanseau P. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47:856–860. doi: 10.1038/ng.3314.
3. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707–713. doi: 10.1038/nature09270.
4. LeBlanc M, Zuber V, Andreassen BK, et al; CARDIoGRAM Consortium. Identifying novel gene variants in coronary artery disease and shared genes with several cardiovascular risk factors. *Circ Res*. 2016;118:83–94. doi: 10.1161/CIRCRESAHA.115.306629.
5. Evans DM, Davey Smith G. Mendelian randomization: new applications in the coming age of hypothesis-free causality. *Annu Rev Genomics Hum Genet*. 2015;16:327–350. doi: 10.1146/annurev-genom-090314-050016.
6. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, Gut I, Lathrop M, Collins R; SEARCH Collaborative Group. SLCO1B1 variants and statin-induced myopathy–a genomewide study. *N Engl J Med*. 2008;359:789–799. doi: 10.1056/NEJMoa0801936.
7. Aminkeng F, Bhavsar AP, Visscher H, et al; Canadian Pharmacogenomics Network for Drug Safety Consortium. A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. *Nat Genet*. 2015;47:1079–1084. doi: 10.1038/ng.3374.
8. Mega JL, Walker JR, Ruff CT, Vandell AG, Nordio F, Deenadayalu N, Murphy SA, Lee J, Mercuri MF, Giugliano RP, Antman EM, Braunwald E, Sabatine MS. Genetics and the clinical response to warfarin and edoxaban: findings from the randomised, double-blind ENGAGE AF-TIMI 48 trial. *Lancet*. 2015;385:2280–2287. doi: 10.1016/S0140-6736(14)61994-2.
9. Do R, Stitziel NO, Won HH, et al; NHLBI Exome Sequencing Project. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*. 2015;518:102–106. doi: 10.1038/nature13917.
10. Helgadottir A, Gretarsdottir S, Thorleifsson G, et al. Variants with large effects on blood lipids and the role of cholesterol and triglycerides in coronary disease. *Nat Genet*. 2016;48:634–639. doi: 10.1038/ng.3561.
11. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet*. 2016;48:709–717. doi: 10.1038/ng.3570.
12. Gaulton KJ, Ferreira T, Lee Y, et al; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat Genet*. 2015;47:1415–1425. doi: 10.1038/ng.3437.
13. Claussnitzer M, Dankel SN, Kim KH, et al. FTO obesity variant circuitry and adipocyte browning in humans. *N Engl J Med*. 2015;373:895–907. doi: 10.1056/NEJMoa1502214.
14. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014;383:1990–1998. doi: 10.1016/S0140-6736(13)62674-4.
15. Beaudoin M, Gupta RM, Won HH, Lo KS, Do R, Henderson CA, Lavoie-St-Amour C, Langlois S, Rivas D, Lehoux S, Kathiresan S, Tardif JC, Musunuru K, Lettre G. Myocardial infarction-associated SNP at 6p24 interferes with MEF2 binding and associates with PHACTR1 expression levels in human coronary arteries. *Arterioscler Thromb Vasc Biol*. 2015;35:1472–1479. doi: 10.1161/ATVBAHA.115.305534.

KEY WORDS: cardiovascular diseases ■ cholesterol ■ genomics ■ lipids ■ precision medicine

A Strategy for Genomic Research on Common Cardiovascular Diseases Aiming at the Realization of Precision Medicine: Personal Insights and Perspectives

Hiroyuki Morita and Issei Komuro

Circ Res. 2016;119:900-903
doi: 10.1161/CIRCRESAHA.116.309802
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7330. Online ISSN: 1524-4571