

## Review Article

# Genome-Wide Association Study of Coronary Artery Disease

**Naomi Ogawa,<sup>1,2</sup> Yasushi Imai,<sup>1,2</sup> Hiroyuki Morita,<sup>1,3</sup> and Ryozo Nagai<sup>1,2</sup>**

<sup>1</sup> Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-8655, Japan

<sup>2</sup> Division of Clinical Genome Informatics, Graduate School of Medicine, The University of Tokyo, Japan

<sup>3</sup> Department of Translational Research for Healthcare and Clinical Science, Graduate School of Medicine, The University of Tokyo, Japan

Correspondence should be addressed to Ryozo Nagai, [nagai-tyk@umin.ac.jp](mailto:nagai-tyk@umin.ac.jp)

Received 31 May 2010; Accepted 25 June 2010

Academic Editor: Tomohiro Katsuya

Copyright © 2010 Naomi Ogawa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Coronary artery disease (CAD) is a multifactorial disease with environmental and genetic determinants. The genetic determinants of CAD have previously been explored by the candidate gene approach. Recently, the data from the International HapMap Project and the development of dense genotyping chips have enabled us to perform genome-wide association studies (GWAS) on a large number of subjects without bias towards any particular candidate genes. In 2007, three chip-based GWAS simultaneously revealed the significant association between common variants on chromosome 9p21 and CAD. This association was replicated among other ethnic groups and also in a meta-analysis. Further investigations have detected several other candidate loci associated with CAD. The chip-based GWAS approach has identified novel and unbiased genetic determinants of CAD and these insights provide the important direction to better understand the pathogenesis of CAD and to develop new and improved preventive measures and treatments for CAD.

## 1. Introduction

Coronary artery disease (CAD) including myocardial infarction (MI) is a leading cause of death worldwide [1, 2]. The well-known conventional coronary risk factors include age, male sex, hypertension, diabetes mellitus, hypercholesterolemia, smoking and family history, which have been repeatedly demonstrated in multiple epidemiological studies [3–5]. Lifestyle and environmental factors play an important role in the pathogenesis; however, genetic predisposition is also thought to contribute to CAD/MI since these diseases cluster in families [6].

In the epidemiological studies using twins, the relative hazard of death among men from CAD when one's twin died of CAD before the age of 55 years, as compared with the hazard when one's twin did not die before 55, was 8.1 for monozygotic twins and 3.8 for dizygotic twins [7]. The recent epidemiological survey in the Framingham study showed that parental cardiovascular disease independently predicted future offspring events [8]. In this survey, participants with

at least one parent with premature CAD had greater risk for events with age-adjusted odds ratios (ORs) of 2.6 for men and 2.3 for women compared with those with no parental CAD. These results support further research into genetic determinants of CAD risk. Elucidating the genetic determinants would improve risk assessment and provide better measures for prevention and treatment.

As the molecular biology and genetics had progressed, the genetic backgrounds were explored in the several genes which were thought to contribute to the pathogenesis of atherosclerosis and conventional coronary risk factors. Candidate gene studies tested the hypothesis that proteins known to be involved in the pathogenesis of atherosclerosis carry variants that affect their protein functions and the risk of developing CAD. In 1992, Cambien et al. [9] explored a possible association between CAD and a variation found in the gene encoding angiotensin-converting enzyme (ACE). The polymorphism ACE/insertion/deletion (ACE/ID) is strongly associated with the level of the circulating enzyme. They reported that the deletion homozygote (DD genotype),

which was associated with higher levels of circulating ACE, is significantly more frequent in patients with MI than in controls.

The representative variants associated with CAD/MI found by candidate gene approach are listed in Table 1. However, those candidate genes were not always reproducible in multiple studies later on. One of the reasons for poor reproducibility is that many of the study samples were not large enough with some exceptions to identify disease-associated genetic variants with odds ratio <2.0. In addition, candidate gene studies only tested a single to few variants for association with CAD and these approaches cannot discover unknown novel variants and also cannot evaluate how strong each variant contribute to the susceptibility to CAD. Therefore, the candidate gene approach resulted in only limited success in the elucidation of genetic risks for CAD.

In parallel with candidate gene studies, other strategies were carried out to interrogate the entire human genome without hypotheses on which genes may be responsible for disease risk. One of the strategies is genome-wide linkage analysis and it is based on the Mendelian cosegregation of a genetic marker within a family. However, great efforts had to be made in order to collect sufficient numbers of affected sibling pairs. Only a small number of studies were successfully performed and few genetic loci (2q21-22, Xq23-26 [10], myocyte enhancer factor-2 (*MEF2A*) [11], arachidonate 5-lipoxygenase-activating protein (*ALOX5AP*) [12], leukotriene A4 hydrolase (*LTA4H*) [13]) were detected to be associated with CAD (Table 2). These genes had never been suggested as causative genes before these family-based studies, suggesting the effectiveness of this approach in detection of novel genetic determinants. However, those associations were not always replicated. Such a family-based study has been frequently used to identify new loci in monogenic diseases, but the application of this strategy to multifactorial diseases is relatively limited.

In this situation, whole genome analysis in a case-control study design had gradually emerged around the beginning of the 21st century. Resources including the single nucleotide polymorphism (SNP) databases, major technological advances in high-throughput genotyping, and methods of data processing and statistical analysis allow researchers to confront limitations in previous approaches. Here, we introduce the representative genome-wide case-control association studies in the next two sections.

## 2. Establishment of J-SNP Database and Whole Genome Approach in Japan

Beginning in 2000, the Prime Minister's Millennium Project (J-SNP) was launched in Japan and about hundred thousand SNPs located in genes or in adjacent regions that might influence the coding sequence of the genes were identified in Japanese population. J-SNP established a web-based database and allowed researchers access to high quality SNP data [14]. Genome-wide association studies using this SNP database were performed in our country in many important

TABLE 1: The well-known genetic polymorphisms which are thought to be associated with myocardial infarction or coronary artery disease.

Location	Gene name/Polymorphisms
17q23	Angiotensin-Converting Enzyme insertion/deletion (intron 16)
1q42-q43	Angiotensinogen Met235Thr, -6G/A
3q21-q25	Angiotensin II type1 Receptor 1166A/C
8q21-q22	Aldosterone Synthase (CYP11B2) -344T/C, Lys173Arg
14q32.1-q32.2	Bradykinin B2 receptor gene -58T/C
6p24.1	Endothelin-1 Lys198Asn
7q36	eNOS Glu298Asp, -786T/C
17q21.32	Glycoprotein IIIa P1A1/A2
5q23-31	Glycoprotein Ia 807T/C
17pter-p12	Glycoprotein Ibα Thr145Met
4q28	βfibrinogen -455G/A
11p11-q12	Prothrombin 20210G/A
7q21.3-q22	PAI-1 4G/5G (promoter region)
7q21.3	Paraoxonase1 Arg192Gln, Leu54Met
8p12-p11.2	Werner Helicase Gene Cys1367Arg
1p36.3	Methylenetetrahydrofolate reductase 677C/T
16q24	NADH/NADPH oxidase p22phox 242C/T, 640A/G
5q31.1	CD14 Monocyte Receptor -260C/T
11q22.3	Stromelysin (MMP3) 5A/6A (promoter region)
20q11.2-q13.1	Gelatinase B (MMP9) -1562C/T
19q13.2	ApolipoproteinE E2/E3/E4
16q21	Cholesteryl Ester Transfer Protein (CETP) Ile405Val
9q31.1	ABCA1 gene Ile823Met
3p25	PPAR-gamma Pro12Ala, Pro115Gln
20q13.11-q13.13	Prostacyclin synthase gene
	The number of 9-bp (CCGCCAGCC) repeats (promoter region)
17q11.2-q21.1	MCP-1 -2518G/A

TABLE 2: Representative loci associated with CAD identified in the family-based studies.

Reporter (year)	Race	locus	LOD score	causative gene
Pajukanta et al. [10]	Finland	2q21-22	3.7	no gene identified
		Xq23-26	3.5	no gene identified
Wang et al. [11]	USA	15q26	4.19	MEF2
Helgadottir et al. [12]	Iceland	13q12-13	2.86	ALOX5AP
Helgadottir et al. [13]	Iceland	17q22	NA	LTA4H

NA = not available.

clinical fields including cardiovascular diseases, diabetes, renal dysfunction and autoimmune collagen diseases.

As the first genome-wide case-control association study in the world, Ozaki et al. [15] used 92,788 gene-based SNP markers and identified that the homozygosity in two SNPs in lymphotoxin A (*LTA*) at 6p21 was significantly associated with increased risk for MI in Japanese (odds ratio (OR) = 1.78). In vitro analyses showed that one functional SNP (Thr26Asn) caused a twofold increase in induction of several inflammation-related cell-adhesion molecules including vascular cell adhesion molecule 1 (VCAM1) in vascular smooth-muscle cells. Moreover, the SNP located in intron 1 of *LTA* enhanced the transcriptional level of *LTA*. These results indicated the variants in the *LTA* are risk factors for MI and implicated *LTA* as a novel pathogenic factor for MI. In the same year, *LTA* knockout mouse was shown to be resistant to atherosclerosis [16]. Double knockout mice of apolipoprotein E (*ApoE*) and *LTA* (*ApoE*<sup>-/-</sup>*LTA*<sup>-/-</sup> mice) showed less extent of atherosclerosis than *ApoE*<sup>-/-</sup>*LTA*<sup>+/+</sup> mice, indicating *LTA* deteriorates atherosclerosis in vivo, consistent with the result of the genetic association study of *LTA* as a genetic risk for atherosclerotic disease.

Subsequently, they identified SNPs that were significantly associated with MI in lectin, galactoside-binding, soluble, 2 gene (*LGALS2*) [17], proteasome subunit alpha type 6 gene (*PSMA6*) [18], myocardial infarction associated transcript (*MIAT*) [19], inter-alpha (globulin) inhibitor 3 gene (*ITIH3*) [20] and BRCA1-associated protein gene (*BRAP*) [21] mainly by functional approaches. All the six causative genetic regions are related to inflammatory process in vasculature and are thought to contribute to the process in atherosclerotic changes through its inflammatory functions and thus increase the risk of MI.

### 3. Chip-Based GWAS and Novel Candidate Genetic Determinants for CAD

**3.1. HapMap Project.** Genome-wide association studies based on the J-SNP database (approximately 100,000 SNPs) have detected several SNPs which had significant association with myocardial infarction. However, the J-SNP database does not cover SNPs in intergene regions. In the meantime, the International HapMap Project was conducted to create a public genome-wide database of common SNPs and enable systematic studies of common SNPs for their potential role in human disease [22, 23]. The Project analyzed DNA samples from 90 people with European ancestry, 90 Yoruba people in Nigeria, 44 Japanese and 45 Han Chinese and has now genotyped over 3.1 million SNPs in each of these populations. However, testing all of these SNPs in a person's chromosomes would be extremely expensive. Adjacent SNPs across the genome are correlated each other, a phenomenon known as linkage disequilibrium and SNPs that are inherited together were compiled into "haplotypes". A haplotype block may contain many SNPs, but only a few "tag" SNPs can provide most of the information on the pattern of genetic variation in the block. The HapMap project identified these "tag" SNPs within haplotypes that uniquely identify these haplotypes. The HapMap data allowed efficient design of Chip-based

genome association studies and allowed investigators to genotype far fewer SNPs while still retaining statistical power to find genetic variants related to common illness.

**3.2. High throughput SNP Genotyping Platforms.** The development of dense genotyping chips enables genotyping up to 1 million SNPs on a single small chip. This chip technology allowed genome-wide association studies (GWAS) to be performed on a large numbers of subjects. Chip-based GWAS typically involves genotyping approximately a few thousands cases with a disease and a few thousands of controls for about 500,000 tag SNPs. Since there are 500,000 comparisons per study, there is a high potential for false positive results. The proposed solution for that is applying the stringent *P* value using the Bonferroni correction for multiple tests. In that case, *P* value will be 0.05 divided by 500,000 and that is 0.0000001 ( $10^{-7}$ ) and this stringent *P* value is often termed as 'genome-wide significance'. The most statistically significant variants identified in the initial case-control analysis are tested for replication in subsequent case-control studies. In GWAS method, associations between SNPs and the diseases are made free of bias of particular candidate genes. This makes the possibility of obtaining novel and unbiased information and provides the important direction to better understand the pathophysiology of the disease. For CAD, three chip-based GWAS were simultaneously reported in 2007 and all of them showed the significant association between CAD and SNPs on chromosome 9p21.

**3.3. 9p21 and other Chromosomal Loci Associated with CAD/MI Detected in Chip-Based GWAS.** Helgadottir et al. enrolled a total of 4587 MI cases and 12,767 controls and genotyped total 305953 SNPs using Illumina Hap300chip (Illumina) [24] (Table 3). All the participants were European descent. They identified disease association variant located in 9p21, adjacent to the tumor suppressor genes cyclin-dependent kinase inhibitor 2A (*CDKN2A*) and cyclin-dependent kinase inhibitor 2B (*CDKN2B*) with great statistical significance. This region had never been estimated to be associated with susceptibility to MI. They showed the allele G of the SNP rs10757278 (Figure 1) showed the strongest association with MI. The ORs for heterozygous and homozygous carriers of the risk allele G were 1.26 and 1.64, respectively. The ORs for early-onset MI (MI before the age of 50 for males and before the age of 60 for females) are 1.49 and 2.02 for heterozygous and homozygous carriers of the risk allele, respectively. They estimated the population attributable risk is 21% for MI in general and 31% for early-onset cases.

The SNPs on chromosome 9p21 associated with MI are located in the same disequilibrium block of the one which contains *CDKN2A* and *CDKN2B*. These genes encode two members of the inhibitors of CDK4 (Ink4) family of cyclin-dependent kinase inhibitors, p16<sup>Ink4a</sup> and p15<sup>Ink4b</sup>, and a completely unrelated protein called ARF. The p16<sup>Ink4a</sup> and p15<sup>Ink4b</sup> which activates retinoblastoma (Rb) family members and ARF which activates p53 were shown to be upregulated in cancer cells. They play a critical role in cell proliferation and aging, senescence and apoptosis [25, 26]. However, sequencing 93 early-onset MI patients across these genes did

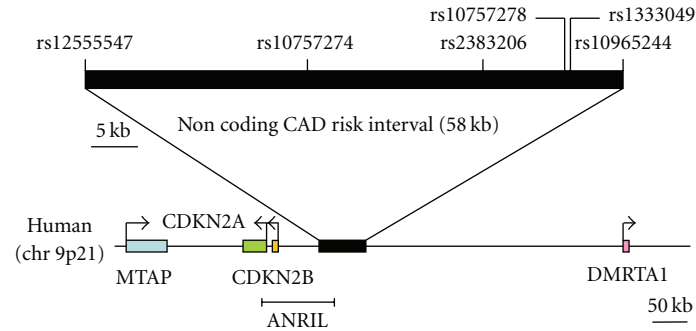


FIGURE 1

FIGURE 1: The 58 kb non-coding CAD risk interval on chromosome 9p21.

not reveal obvious causal functional variants or variants that could account for the correlation of rs10757278 to MI. The linkage disequilibrium block also contains two exons of the transcript hypothetical methylthioadenosine phosphorylase fusion protein mRNA, however the functional significance of the variants in this region remains to be elucidated.

McPherson et al. also identified a 58-kilobase region on chromosome 9p21 that was consistently associated with CAD in six independent samples consisted of 4306 cases and 20119 controls from four Caucasian populations [27]. They identified two SNPs rs10757274 and rs2383206 on 9p21 that are significantly associated with incident of CAD. The risk allele was associated with and ~15 to 20% increase risk of CAD in the 50% individuals who are heterozygous and ~30 to 40% increase in the 25% who are homozygous for the allele.

They further genotyped surrounding region of these SNPs in detail and found that eight additional SNPs at the locus spanning a 58-kb region were significantly associated with CAD. Again, the 58-kb region does not contain any annotated genes. however, the region overlaps a newly annotated noncoding RNA called noncoding RNA in the *INK* locus (*ANRIL*) [28]. *ANRIL* consists of 20 exons subjected to alternative splicing. The whole blood RNA expression levels of short variants of *ANRIL* are increased and the expression levels of the long variant are decreased in subjects homozygous for the risk alleles. There is also a positive correlation between transcript levels of the long variant of *ANRIL* and *CDKN2B* [29].

The study using genetically engineered mice showed that the deletion of orthologous 70-kb non-coding interval on mouse chromosome 4 (chr4<sup>Δ70kb/Δ70kb</sup> mice) highly reduced cardiac expression of two neighboring genes *CDKN2A* and *CDKN2B*. Primary culture of smooth muscle cells from chr4<sup>Δ70kb/Δ70kb</sup> mice showed increased proliferation and diminished senescence, the features relevant to atherosclerosis [30]. These findings indicated the noncoding interval is involved in the disease process via gene-regulatory effects on *CDKN2A* and *CDKN2B*.

The Wellcome Trust Case Control Consortium (WTCCC) study which enrolled 1926 case subjects with CAD and 2938 controls also reported the powerful association

between the SNPs on chromosome 9p21 and CAD [31]. The strongest signal was seen at rs1333049, however the associations were seen for SNPs across >100 kb. Then, they further looked for replication in the German MI family study which involved 875 MI cases and 1644 controls using the GeneChip Human Mapping 500K Array Set (Affymetrix) [32]. The Same locus on chromosome 9p21 (rs1333049) had the strongest association with CAD in both studies with the risk increased by 36% per copy of the C allele. Of the nine loci which were shown to be strongly associated with CAD, two of these loci were able to replicate in the German MI study: chromosome 6q25.1 (rs6922269) and chromosome 2q36.3 (rs2943634). Further, the combined analysis of the two studies revealed four additional loci significantly associated with CAD: chromosome 1p13.3 (rs599839), 1q41 (rs17465637), 10q11.21 (rs501120) and 15q22.33 (rs17228212). In 2009, they performed another replication study with 11550 cases with CAD and 11205 controls from 9 European studies [33]. Other than the 9p21 locus, they confirmed significant association at 1p13.3 (rs599839), 1q41 (rs3008621) and 10q11.21 (rs501120). They were not able to show the significant association with 6q25.1 and 2q36.3 and there was no evidence for association with the locus at 15q22.33. The four loci (9p21, 1p13.3, 1q41 and 10q11.21) act independently and cumulatively increased the risk for CAD by 15% per additional risk allele. The genes located within or adjacent to the these four loci are listed in Table 4. The locus at chromosome 1p13.3 has been shown to be associated with increased plasma LDL cholesterol, and thus may contribute to CAD development [34–37].

The locus at 10q11.21 lies adjacent to the chemokine (C-X-C motif) ligand 12 gene (*CXCL12*) which encodes stromal cell-derived factor-1, a chemokine which plays a important role in stem-cell homing and regeneration of myocardial tissue in ischemic cardiomyopathy [38] and in promoting angiogenesis by recruiting endothelial progenitor cells from the bone marrow [39]. The SNPs at 1q41 locates within the melanoma inhibitory activity family, member 3 (*MIA3*) gene [40]. Underlying mechanism how these genetic loci affect the pathogenesis of CAD need to be further investigated.

The meta-analysis of aforementioned three studies [24, 27, 32] and 7 additional case-control studies successfully

TABLE 3: Representatives of chip-based GWAS of CAD/MI.

Author, year	Phenotype	No. of cases/controls	Chromosomal loci	OR
Helgadottir et al. [24]	MI	4587/12767	9p21	1.28
McPherson et al. [27]	CAD	3505/18745	9p21	1.26 (CCHS study)
			9p21	1.16 (CCHS study)
WTCCC [31]	CAD	2000/3000	9p21.3	1.47
Samani et al. [32]	CAD	2801/4582	9p21.3	1.28 (adjusted German)
			6q25.1	1.23 (adjusted German)
			2q36.3	1.08 (adjusted German)
			1p13.3	1.29
			1q41	1.2
			10q11.21	1.33
			15q22.33	1.21
CAD Consortium, [33]	CAD	11550/11205	9p21	1.2
			1p13.3	1.13
			1p41	1.1
			10q11.21	1.11
Clarke et al. [59]	CAD	7991/7946	6q26-27	1.51

TABLE 4: Genes located within or adjacent to the loci associated with CAD/MI.

Chromosome loci	Genes
1p33	PSRC1, CELRS2, MYBPHL, SORT1
1q41	MIA3
2q33	WDR12
2q36.3	no recognized genes
3q22.3	MRAS
6p24	PHACTR1
6q25.1	MTHFD1L
6q26-27	LPA
9p21	p16/CDKN2A, p15/CDKN2B, p14/ARF, MTAP, ANRIL
10q11.21	CXCL12
12q24	3SH2B3
21q22	MRPS6, SLC5A3, KCNE2

replicated the significant association between the risk allele (C) of the lead SNP, rs1333049 at chromosome 9p21 and risk of CAD (OR = 1.29) [41]. These study have analyzed primarily on European descent, however, since the allele frequency differs among the different ethnic groups, the risk of CAD related to the SNPs at 9p21 may differ among each ethnic group. Since then, the replicated results in other ethnics such as Chinese, Japanese and Pakistanis are published and the association of the SNPs at 9p21 with CAD seems to be consistent among various ethnic groups [42–45].

It is noteworthy that the SNPs in 9p21 region are also found to be associated with variety of diseases such as ischemic stroke (OR = 1.01–1.21) [46, 47], abdominal aortic aneurysm (OR = 1.31), intracranial aneurysm (OR = 1.29)

[48], peripheral artery disease (OR = 1.29) [49], incident heart failure (OR = 1.17) [50], perioperative myocardial injury after coronary artery bypass graft surgery [51], type 2 diabetes (OR = 1.20) [52, 53].

More recently, the loci other than 9p21 have been shown to be associated with CAD or MI. Erdmann et al. applied less stringent statistical thresholds on their GWAS for CAD to identify any dismissed SNPs with modest effects or low allele frequencies and they found one new locus on 3q22.3 in muscle RAS oncogene homolog (*MRAS*) (OR = 1.15), the gene thought to play an important role in inflammation [54, 55].

There is another GWAS which identified the SNPs in the gene related to inflammation to be associated with MI [56]. They found five SNPs which affect eosinophil counts in blood in Icelandic population and reported that a nonsynonymous SNP at 12q24 in *SH2B* adaptor protein gene (*3SH2B3*) was associated with MI significantly (OR = 1.13).

The GWAS of early-onset MI revealed 9 loci which have significant association [57]. Three of them were newly identified in the study: (i) an intergenic region between *MRPS6* (mitochondrial ribosomal protein S6), *SLC5A3* (solute carrier family 5 (sodium/myo-inositol cotransporter), member 3) and *KCNE2* (potassium voltage-gated channel, Isk-related family, member 2) on chromosome 21q22 (OR = 1.19), 6p24 in *PHACTR1* (phosphatase and actin regulator 1) (OR = 1.13) and 2q33 in *WDR12* (WD repeat domain 12) (OR = 1.17). The mechanism by which genes at these three regions increases the risk of MI needs to be elucidated. In addition to the common variant, copy number variations can be analyzed by SNP chip and there were no common or rare copy number variations associated with risk of early-onset MI in this study.

The group from the WTCCC/German MI study further conducted a genome-wide haplotype association study for

the first time and identified the *SLC22A3-LPAL2-LPA* gene cluster on 6q26-27 as a strong susceptibility locus for CAD (OR = 1.8) [58]. An increased level of Lp(a) lipoprotein is a classical hereditary risk for CAD. Clarke et al. identified three chromosomal regions (6q26-27, 9p21 and 1p13) were strongly associated with the risk of CAD using the Human CVD Bead Chip which included 48742 markers relevant to cardiovascular disease on 6500 subjects. Among them, the *LPA* locus on 6q26-27 encoding Lp(a) lipoprotein had the strongest association. They identified two *LPA* variants that were strongly associated with both an increased level of Lp(a) lipoprotein and an increased risk of CAD (OR = 1.70 and 1.92) [59]. Both variants were strongly associated with a reduced copy number in *LPA* kringle IV-type 2 repeats and an increased level of Lp(a) lipoprotein. After adjustment for the Lp(a) lipoprotein level, the association between the *LPA* genotype score and the risk of CAD was abolished. The importance of classical risk factor Lp(a) was reemphasized by the GWAS.

#### 4. Limitation of GWAS

Initially, GWAS have been primarily assessed only on European descent and the results of these GWAS may not be applicable to other ethnics due to wide difference of distribution of SNPs and allele frequency. Further studies for various ethnicity need to be done with use of newer chips which contains 1 million SNPs to increase coverage.

The CAD associated loci have been found in regions without known gene-encoding loci. Therefore, further studies will be required to elucidate the exact functional mechanism by which these loci modulate CAD risk.

Utility of genotyping 9p21 for clinical risk assessment is controversial [60–62]. The odds ratios for CAD risk in each selected SNPs are small (around 1.2) and explain only a small proportion of the heritable, genetic component of susceptibility to the disease. Newer susceptibility loci for CAD need to be validated with replication studies and in the future, we should evaluate the genetic risks by combining multiple independent common variants susceptible for CAD.

The GWAS method is supported by the common disease-common variant hypothesis, which predicts that genetic variants causing common disease exist frequently, but each variant only have a small effect on disease susceptibility. Another hypothesis is the rare variant hypothesis. Rare variants have a minor allele frequency of less than 1%. The rare variant hypothesis postulates that common disease is caused by multiple rare variants which have a strong causative effect on disease and this hypothesis was confirmed in colorectal adenomas [63, 64]. Rare variants cannot be captured by GWAS and requires whole genome sequencing using next generation sequencing system.

In addition, other types of variants, such as insertion-deletion variant, block substitution and inversion variant, so called structural variants may account for important contributors to the diseases and are also hard to detect by the chip-based method. The next generation sequencing method is also helpful to find these structural variants.

#### 5. Conclusion

- (1) The SNPs Data from the HapMap project and development of new chip technology enabled genotyping large amount of common variants simultaneously and contributed to efficiently identify gene loci affecting susceptibility to common diseases including CAD.
- (2) The region at 9p21 was shown to be significantly associated with CAD in 2007 and comprehensive replication across multiple studies provides unequivocal evidence that this locus is associated with CAD in European descent. This region is also associated with abdominal aneurysm, intracranial aneurysm and type 2 diabetes, and seems to be a very important region for various diseases.
- (3) Since the odds ratios of the risk allele at 9p21 for CAD are small, screening for this risk allele probably affects little, if any, to the each individual's risk prediction. Using genomic tests to improve existing risk models would likely require combining the effects of multiple common genetic variants.
- (4) Rare variants and structural variants which cannot be captured by GWAS need to be searched by whole genome sequencing.
- (5) GWAS approach has identified novel and unbiased genetic contributors to CAD and these insights provide the important direction to better understand the pathogenesis of CAD and to develop new and improved preventive measures and treatments for CAD.

#### References

- [1] A. D. Lopez, C. D. Mathers, M. Ezzati, D. T. Jamison, and C. J. Murray, "Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data," *The Lancet*, vol. 367, no. 9524, pp. 1747–1757, 2006.
- [2] D. Lloyd-Jones, R. J. Adams, T. M. Brown et al., "Heart disease and stroke statistics—2010 update: a report from the American heart association," *Circulation*, vol. 121, no. 7, pp. 948–954, 2010.
- [3] T. R. Dawber, F. E. Moore, and G. V. Mann, "Coronary heart disease in the Framingham study," *American Journal of Public Health and the Nation's Health*, vol. 47, pp. 4–24, 1957.
- [4] P. W. F. Wilson, "Established risk factors and coronary artery disease: the Framingham Study," *American Journal of Hypertension*, vol. 7, no. 7, pp. 7S–12S, 1994.
- [5] R. Nagai, "Current status of the background of patients with coronary artery disease in Japan—the Japanese coronary artery disease study (The JCAD Study)," *Circulation Journal*, vol. 70, no. 10, pp. 1256–1262, 2006.
- [6] S. Shea, R. Ottman, and C. Gabrieli, "Family history as an independent risk factor for coronary artery disease," *Journal of the American College of Cardiology*, vol. 4, no. 4, pp. 793–801, 1984.

- [7] M. E. Marenberg, N. Risch, L. F. Berkman, B. Floderus, and U. De Faire, "Genetic susceptibility to death from coronary heart disease in a study of twins," *The New England Journal of Medicine*, vol. 330, no. 15, pp. 1041–1046, 1994.
- [8] D. M. Lloyd-Jones, B.-H. Nam, R. B. D'Agostino Sr. et al., "Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring," *Journal of the American Medical Association*, vol. 291, no. 18, pp. 2204–2211, 2004.
- [9] F. Cambien, O. Poirier, L. Lecerf et al., "Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction," *Nature*, vol. 359, no. 6396, pp. 641–644, 1992.
- [10] P. Pajukanta, M. Cargill, L. Viitanen, et al., "Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland," *The American Journal of Human Genetics*, vol. 67, no. 6, pp. 1481–1493, 2000.
- [11] L. Wang, C. Fan, S. E. Topol, E. J. Topol, and Q. Wang, "Mutation of MEF2A in an inherited disorder with features of coronary artery disease," *Science*, vol. 302, no. 5650, pp. 1578–1581, 2003.
- [12] A. Helgadottir, A. Manolescu, G. Thorleifsson et al., "The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke," *Nature Genetics*, vol. 36, no. 3, pp. 233–239, 2004.
- [13] A. Helgadottir, A. Manolescu, A. Helgason et al., "A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction," *Nature Genetics*, vol. 38, no. 1, pp. 68–74, 2006.
- [14] M. Hirakawa, T. Tanaka, Y. Hashimoto, M. Kuroda, T. Takagi, and Y. Nakamura, "JSNP: a database of common gene variations in the Japanese population," *Nucleic Acids Research*, vol. 30, no. 1, pp. 158–162, 2002.
- [15] K. Ozaki, Y. Ohnishi, A. Iida et al., "Functional SNPs in the lymphotoxin- $\alpha$  gene that are associated with susceptibility to myocardial infarction," *Nature Genetics*, vol. 32, no. 4, pp. 650–654, 2002.
- [16] S. A. Schreyer, C. M. Vick, and R. C. Leboeuf, "Loss of lymphotoxin- $\alpha$  but not tumor necrosis factor- $\alpha$  reduces atherosclerosis in mice," *The Journal of Biological Chemistry*, vol. 277, no. 14, pp. 12364–12368, 2002.
- [17] K. Ozaki, K. Inoue, H. Sato et al., "Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin- $\alpha$  secretion in vitro," *Nature*, vol. 429, no. 6987, pp. 72–75, 2004.
- [18] K. Ozaki, H. Sato, A. Iida et al., "A functional SNP in PSMA6 confers risk of myocardial infarction in the Japanese population," *Nature Genetics*, vol. 38, no. 8, pp. 921–925, 2006.
- [19] N. Ishii, K. Ozaki, H. Sato et al., "Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction," *Journal of Human Genetics*, vol. 51, no. 12, pp. 1087–1099, 2006.
- [20] Y. Ebana, K. Ozaki, K. Inoue et al., "A functional SNP in ITIH3 is associated with susceptibility to myocardial infarction," *Journal of Human Genetics*, vol. 52, no. 3, pp. 220–229, 2007.
- [21] K. Ozaki, H. Sato, K. Inoue et al., "SNPs in BRAP associated with risk of myocardial infarction in Asian populations," *Nature Genetics*, vol. 41, no. 3, pp. 329–333, 2009.
- [22] International HapMap Consortium, "A haplotype map of the human genome," *Nature*, vol. 437, no. 7063, pp. 1299–1320, 2005.
- [23] K. A. Frazer, D. G. Ballinger, D. R. Cox et al., "A second generation human haplotype map of over 3.1 million SNPs," *Nature*, vol. 449, no. 7164, pp. 851–861, 2007.
- [24] A. Helgadottir, G. Thorleifsson, A. Manolescu et al., "A common variant on chromosome 9p21 affects the risk of myocardial infarction," *Science*, vol. 316, no. 5830, pp. 1491–1493, 2007.
- [25] W. Y. Kim and N. E. Sharpless, "The regulation of INK4/ARF in cancer and aging," *Cell*, vol. 127, no. 2, pp. 265–275, 2006.
- [26] M. Collado and M. Serrano, "Senescence in tumours: evidence from mice and humans," *Nature Reviews Cancer*, vol. 10, no. 1, pp. 51–57, 2010.
- [27] R. McPherson, A. Pertsemlidis, N. Kavaslar et al., "A common allele on chromosome 9 associated with coronary heart disease," *Science*, vol. 316, no. 5830, pp. 1488–1491, 2007.
- [28] E. Pasmant, I. Laurendeau, D. Héron, M. Vidaud, D. Vidaud, and I. Bièche, "Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF," *Cancer Research*, vol. 67, no. 8, pp. 3963–3969, 2007.
- [29] O. Jarinova, A. F. R. Stewart, R. Roberts et al., "Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 29, no. 10, pp. 1671–1677, 2009.
- [30] A. Visel, Y. Zhu, D. May et al., "Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice," *Nature*, vol. 464, no. 7287, pp. 409–412, 2010.
- [31] P. R. Burton, D. G. Clayton, L. R. Cardon et al., "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls," *Nature*, vol. 447, no. 7145, pp. 661–678, 2007.
- [32] N. J. Samani, J. Erdmann, A. S. Hall et al., "Genomewide association analysis of coronary artery disease," *The New England Journal of Medicine*, vol. 357, no. 5, pp. 443–453, 2007.
- [33] N. J. Samani, P. Deloukas, and J. Erdmann, "Large scale association analysis of novel genetic loci for coronary artery disease," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 29, no. 5, pp. 774–780, 2009.
- [34] C. Wallace, S. J. Newhouse, P. Braund et al., "Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia," *American Journal of Human Genetics*, vol. 82, no. 1, pp. 139–149, 2008.
- [35] M. S. Sandhu, D. M. Waterworth, S. L. Debenham et al., "LDL-cholesterol concentrations: a genome-wide association study," *The Lancet*, vol. 371, no. 9611, pp. 483–491, 2008.
- [36] N. J. Samani, P. S. Braund, J. Erdmann et al., "The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol," *Journal of Molecular Medicine*, vol. 86, no. 11, pp. 1233–1241, 2008.
- [37] M. E. Keebler, C. L. Sanders, A. Surti, C. Guiducci, N. P. Burt, and S. Kathiresan, "Association of blood lipids with common DNA sequence variants at 19 genetic loci in the multiethnic United States national health and nutrition examination survey III," *Circulation*, vol. 120, no. 3, pp. 238–243, 2009.
- [38] A. T. Askari, S. Unzek, Z. B. Popovic et al., "Effect of stromal-cell-derived factor 1 on stem-cell homing and tissue regeneration in ischaemic cardiomyopathy," *The Lancet*, vol. 362, no. 9385, pp. 697–703, 2003.

- [39] H. Zheng, G. Fu, T. Dai, and H. Huang, "Migration of endothelial progenitor cells mediated by stromal cell-derived factor-1 $\alpha$ /CXCR4 via PI3K/Akt/eNOS signal transduction pathway," *Journal of Cardiovascular Pharmacology*, vol. 50, no. 3, pp. 274–280, 2007.
- [40] A.-K. Bosserhoff and R. Buettner, "Expression, function and clinical relevance of MIA (melanoma inhibitory activity)," *Histology and Histopathology*, vol. 17, no. 1, pp. 289–300, 2002.
- [41] H. Schunkert, A. Götz, P. Braund et al., "Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease," *Circulation*, vol. 117, no. 13, pp. 1675–1684, 2008.
- [42] H. Ding, Y. Xu, X. Wang et al., "9p21 is a shared susceptibility locus strongly for coronary artery disease and weakly for ischemic stroke in Chinese han population," *Circulation*, vol. 120, no. 4, pp. 338–346, 2009.
- [43] A. Maitra, D. Dash, S. John et al., "A common variant in chromosome 9p21 associated with coronary artery disease in Asian Indians," *Journal of Genetics*, vol. 88, no. 1, pp. 113–118, 2009.
- [44] Y. Hiura, Y. Fukushima, M. Yuno et al., "Validation of the association of genetic variants on chromosome 9p21 and 1q41 With myocardial infarction in a Japanese population," *Circulation Journal*, vol. 72, no. 8, pp. 1213–1217, 2008.
- [45] D. Saleheen, M. Alexander, A. Rasheed, et al., "Association of the 9p21.3 locus with risk of first-ever myocardial infarction in pakistanis," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 30, no. 7, pp. 1467–1473, 2010.
- [46] A. Gschwendtner, S. Bevan, J. W. Cole et al., "Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke," *Annals of Neurology*, vol. 65, no. 5, pp. 531–539, 2009.
- [47] M. Matarin, W. M. Brown, A. Singleton, J. A. Hardy, and J. F. Meschia, "Whole genome analyses suggest ischemic stroke and heart disease share an association with polymorphisms on chromosome 9p21," *Stroke*, vol. 39, no. 5, pp. 1586–1589, 2008.
- [48] A. Helgadottir, G. Thorleifsson, K. P. Magnusson et al., "The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm," *Nature Genetics*, vol. 40, no. 2, pp. 217–224, 2008.
- [49] C. Cluett, M. M. McDermott, J. Guralnik et al., "The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people," *Circulation*, vol. 120, no. 4, pp. 347–353, 2009.
- [50] K. Yamagishi, A. R. Folsom, W. D. Rosamond, and E. Boerwinkle, "A genetic variant on chromosome 9p21 and incident heart failure in the ARIC study," *European Heart Journal*, vol. 30, no. 10, pp. 1222–1228, 2009.
- [51] K.-Y. Liu, J. D. Muehlschlegel, T. E. Perry et al., "Common genetic variants on chromosome 9p21 predict perioperative myocardial injury after coronary artery bypass graft surgery," *Journal of Thoracic and Cardiovascular Surgery*, vol. 139, no. 2, pp. 483–488.e2, 2010.
- [52] E. Zeggini, M. N. Weedon, C. M. Lindgren et al., "Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes," *Science*, vol. 316, no. 5829, pp. 1336–1341, 2007.
- [53] L. J. Scott, K. L. Mohlke, L. L. Bonnycastle et al., "A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants," *Science*, vol. 316, no. 5829, pp. 1341–1345, 2007.
- [54] J. Erdmann, A. Großhennig, P. S. Braund et al., "New susceptibility locus for coronary artery disease on chromosome 3q22.3," *Nature Genetics*, vol. 41, no. 3, pp. 280–282, 2009.
- [55] Y. Yoshikawa, T. Satoh, T. Tamura et al., "The M-Ras-RA-GEF-2-Rap1 pathway mediates tumor necrosis factor- $\alpha$ -dependent regulation of integrin activation in splenocytes," *Molecular Biology of the Cell*, vol. 18, no. 8, pp. 2949–2959, 2007.
- [56] D. F. Gudbjartsson, U. S. Bjornsdottir, E. Halapi et al., "Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction," *Nature Genetics*, vol. 41, no. 3, pp. 342–347, 2009.
- [57] S. Kathiresan, D. Ardisino, P. M. Mannucci et al., "Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants," *Nature Genetics*, vol. 41, no. 3, pp. 334–341, 2009.
- [58] D.-A. Trégouët, I. R. König, J. Erdmann et al., "Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease," *Nature Genetics*, vol. 41, no. 3, pp. 283–285, 2009.
- [59] R. Clarke, J. F. Peden, J. C. Hopewell et al., "Genetic variants associated with Lp(a) lipoprotein level and coronary disease," *The New England Journal of Medicine*, vol. 361, no. 26, pp. 2518–2528, 2009.
- [60] A. Brautbar, C. M. Ballantyne, K. Lawson et al., "Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the atherosclerosis risk in communities study," *Circulation*, vol. 120, no. 3, pp. 279–285, 2009.
- [61] K. L. Ellis, A. P. Pilbrow, C. M. Frampton, et al., "A common variant at chromosome 9P21.3 is associated with age of onset of coronary disease but not subsequent mortality," *Circulation Cardiovascular Genetics*, vol. 3, no. 3, pp. 283–293, 2010.
- [62] N. P. Paynter, D. I. Chasman, J. E. Buring, D. Shiffman, N. R. Cook, and P. M. Ridker, "Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3," *Annals of Internal Medicine*, vol. 150, no. 2, pp. 65–72, 2009.
- [63] W. Bodmer and C. Bonilla, "Common and rare variants in multifactorial susceptibility to common diseases," *Nature Genetics*, vol. 40, no. 6, pp. 695–701, 2008.
- [64] N. S. Fearnhead, J. L. Wilding, B. Winney et al., "Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 45, pp. 15992–15997, 2004.