

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/31617> holds various files of this Leiden University dissertation.

Author: Postmus, Iris

Title: Genetics, statins, and lipid metabolism in cardiovascular disease

Issue Date: 2015-01-28

**Replication of LDL GWAS hits in PROSPER/PHASE as
validation for future (pharmaco)genetic analyses**

Stella Trompet, Anton JM de Craen, Iris Postmus, Ian Ford, Naveed Sattar,
Muriel Caslake, David J Stott, Brendan M Buckley, Frank Sacks,
James J Devlin, P Eline Slagboom, Rudi GJ Westendorp, J Wouter Jukema
on behalf of the PROSPER study group

BMC Med Genet. 2011;12:131

Abstract

The PHarmacogenetic study of Statins in the Elderly at risk (PHASE) is a genome wide association study in the PROspective Study of Pravastatin in the Elderly at risk for vascular disease (PROSPER) that investigates the genetic variation responsible for the individual variation in drug response to pravastatin. Statins lower LDL cholesterol in general by 30%, however not in all subjects. Moreover, clinical response is highly variable and adverse effects occur in a minority of patients. In this report we first describe the rationale of the PROSPER/PHASE project and second show that the PROSPER/PHASE study can be used to study pharmacogenetics in the elderly. The genome wide association study (GWAS) was conducted using the Illumina 660K-Quad beadchips following manufacturer's instructions. After a stringent quality control 557,192 SNPs in 5,244 subjects were available for analysis. To maximize the availability of genetic data and coverage of the genome, imputation up to 2.5 million autosomal CEPH HapMap SNPs was performed with MACH imputation software. The GWAS for LDL cholesterol is assessed with an additive linear regression model in PROBABEL software, adjusted for age, sex, and country of origin to account for population stratification. Forty-two SNPs reached the GWAS significant threshold of $p=5.0e-08$ in 5 genomic loci (*APOE/APOC1*; *LDLR*; *FADS2/FEN1*; *HMGCR*; *PSRC1/CELSR5*). The top SNP (rs445925, chromosome 19) with a p-value of $p=2.8e-30$ is located within the *APOC1* gene and near the *APOE* gene. The second top SNP (rs6511720, chromosome 19) with a p-value of $p=5.22e-15$ is located within the *LDLR* gene. All 5 genomic loci were previously associated with LDL cholesterol levels, no novel loci were identified. Replication in WOSCOPS and CARE confirmed our results. With the GWAS in the PROSPER/PHASE study we confirm the previously found genetic associations with LDL cholesterol levels. With this proof-of-principle study we show that the PROSPER/PHASE study can be used to investigate genetic associations in a similar way to population based studies. The next step of the PROSPER/PHASE study is to identify the genetic variation responsible for the variation in LDL cholesterol lowering in response to statin treatment in collaboration with other large trials.

Introduction

Cardiovascular disease is the leading cause of death in industrialized countries at old age. Advancing age is one of the most important risk factors for cardiovascular disease¹. With the rising number of elderly people in our society cardiovascular disease has a major impact on healthcare². The prevention of cardiovascular disease is critically dependent on lipid lowering therapy including the 3-hydroxymethyl-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors (statins). Statins are the most prescribed class of drugs worldwide and therapy is generally associated with a reduction of cardiovascular events by 20-30%. However, clinical response is highly variable and adverse effects occur in a minority of patients³. Recent research provides evidence that genetic variation contributes importantly to this variable drug response⁴.

Pharmacogenomics focuses on unraveling the genetic determinants of such variable drug responses, both in intended, beneficial effects and unintended, adverse effects⁵. Therefore, we here present the PHarmacogenetic study of Statin in the Elderly at risk (PHASE) a genome wide association study (GWAS) in the PROspective Study of Pravastatin in the Elderly at Risk for vascular disease (PROSPER)⁶ investigating the genetic variation responsible for the individual variation in drug response funded by the European Union's Seventh Framework Programme. To validate the GWAS performed in the PHASE study, we executed a proof-of-principle study to investigate the underlying genetic variation in LDL cholesterol levels.

Recent GWA studies have identified several new loci that influence circulating levels of blood lipids with around 95 loci showing statistical associations with circulating total cholesterol levels, HDL cholesterol, LDL cholesterol, and triglycerides⁷. These GWA studies are executed in population based studies with various age groups, however the elderly (age >75 years) are rarely represented in these studies. With this proof-of-principle study we provide a testing frame to show that the PROSPER/PHASE study has sufficient statistical power to find genome wide statistical significant associations in quantitative traits such as LDL cholesterol in an elderly population. We replicated our findings from the PROSPER/PHASE study in two independent cohorts to validate that our results contain no false positive findings.

Methods

Study population

PROSPER was an investigator-driven, prospective multi-national randomized placebo-controlled trial to assess whether treatment with pravastatin diminishes the risk of major vascular events in the elderly^{6,8}. Between December 1997 and May 1999, we screened and enrolled subjects in Scotland, Ireland, and the Netherlands. Men and women aged 70-82 years were recruited if they had pre-existing vascular disease or were at increased risk of such disease because of smoking, hypertension, or diabetes. A total number of 5804 subjects, of whom more than 50% was female, were randomly assigned to pravastatin or placebo. Various clinical laboratory measurements were carried out like inflammatory markers (CRP and various cytokines) and other biochemical substrates (e.g. glucose, leptin) at baseline and during follow-up. The protocol of the PROSPER study meets the criteria of the Declaration of Helsinki and was approved by the Medical Ethics Committees of each participating institution. Written informed consent was obtained from all participating subjects.

LDL cholesterol

Plasma lipids and lipoproteins were measured twice during the screening phase, i.e. at the beginning and end of the single-blind, placebo "run-in" phase according to the standardized Lipid Research Clinics protocol. Baseline LDL cholesterol levels were taken as the average of these 2 determinations prior to randomization to statin treatment. Total cholesterol (TC), HDL cholesterol, and triglycerides were assessed after an overnight fast, LDL cholesterol was calculated by the Friedewald formula, as previously described⁸.

Genotyping

The genotyping was conducted using the Illumina 660-Quad beadchips following manufacturer's instructions. These beadchips contain 657,366 single nucleotide polymorphism (SNP) and copy number variants (CNV) probes. After genotyping, samples and genetic markers were subjected to a stringent quality control protocol. From the 5763 samples with DNA available that underwent genotyping, 519 samples (9%) were excluded during the quality control (figure 1). Excluded were 18 duplicated samples, 219 samples with a call rate <97.5%, 11 samples with an excess for heterozygosity, 40 samples of non-caucasian origin, 170 samples with familial relationships (IBD>0.35), and 61 samples with a gender mismatch. From the 657,366 probes on the beadchips, 95,876 probes were filtered based on CNV intensity. Moreover, 4,298 SNPs were excluded with a call rate < 95%, leaving us with 557,192

SNPs for analysis. To maximize the availability of genetic data and coverage of the genome, imputation up to 2.5 million autosomal CEPH HapMap SNPs was performed with MACH imputation software based on the HapMap built II release 23. To assess accuracy of the imputed genotypes, we compared the imputation output with SNPs that had been previously genotyped on other platforms.

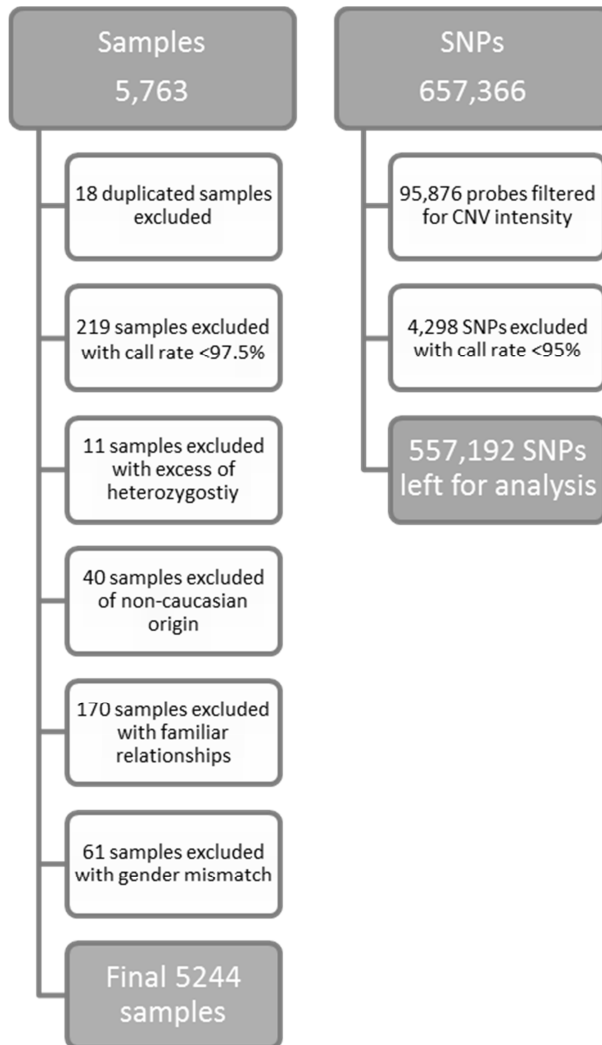


Figure 1. Flow chart of the Quality Control of the PROSPER/PHASE study

Statistical Analysis

Genome wide association analysis was performed with PROBABEL software specialized in genetic association analysis with imputed data taking the probability of the genotype into account (<http://www.genabel.org/>). With analyzing imputed genotypes, the observed allele count is replaced by the imputation's estimated dosage. For the continuous trait, baseline LDL cholesterol levels, an additive linear regression model was used to assess estimates and standard errors. The model was adjusted for sex and age, and country to correct for the within-study population structure. Standard errors for the regression estimates were calculated with model-robust methods. The analysis of 2.5 million SNPs at once poses a multiple testing problem. After the use of a Bonferroni correction, the threshold for genome wide significant results was set at $5.0e-08$.

Replication

Associations with a genome-wide significant p-value of $5.0e-08$ were replicated in two independent cohorts, the West of Scotland Coronary Prevention Study (WOSCOPS)⁹ and the Cholesterol and Recurrent Events (CARE) trial¹⁰. The WOSCOPS study was a double blind randomized placebo-controlled clinical trial in which 6595 men (age range 45-64 years) with hypercholesterolemia and no history of myocardial infarction were treated with 40 mg pravastatin (N=3302) or placebo (N=3293). GWAS data and baseline LDL cholesterol levels were available for 431 subjects. The CARE study was a double blind randomized placebo-controlled clinical trial in which 4159 patients (age range 21-75 years) were treated with 40 mg pravastatin (N=2081) or placebo (N=2078). GWAS data and baseline LDL cholesterol levels were available for 751 subjects. The significance level for the replication SNPs was set at p-value < 0.05.

Results

Table 1 shows the baseline characteristics of the subjects participating in the PROSPER and the PROSPER/PHASE study. This table shows that the genotyped subjects in the PROSPER/PHASE study are representative of the total study population of the PROSPER study, since no major discrepancies exist between the two study sets. The mean age of all subjects at study entry was 75.3 years and about 50% of the participants were female.

In Figure 2 the QQ-plot of the genome-wide association study with baseline LDL levels within the PROSPER/PHASE study is shown. In this plot it is shown that no genomic inflation has occurred in this analyses ($\lambda=1.077$) and that population stratification is sufficiently controlled for. In figure 3 the results of the genome-wide

Table 1. Baseline characteristics of the PROSPER/PHASE study

	PROSPER study (n=5804)	PROSPER/PHASE study (n=5244)
Continuous variables (mean, SD)		
Age (years)	75.3 (3.3)	75.3 (3.4)
Education (years)	15.1 (2.0)	15.1 (2.0)
Systolic blood pressure (mmHg)	154.7 (21.8)	154.6 (21.9)
Diastolic blood pressure (mmHg)	83.8 (11.5)	83.7 (11.4)
Height (cm)	165.2 (9.4)	165.2 (9.4)
Weight (kg)	73.4 (13.4)	73.3 (13.4)
Body mass index (kg/m ²)	26.8 (4.2)	26.8 (4.2)
Total cholesterol (mmol/L)	5.7 (0.9)	5.7 (0.9)
LDL cholesterol (mmol/L)	3.8 (0.8)	3.8 (0.8)
HDL cholesterol (mmol/L)	1.3 (0.3)	1.3 (0.4)
Triglycerides (mmol/L)	1.5 (0.7)	1.5 (0.7)
Categorical variables (n, %)		
Males	2804 (48.3)	2524 (48.1)
Current smoker	1558 (26.8)	1392 (26.5)
History of diabetes	623 (10.7)	544 (10.4)
History of hypertension	3592 (61.9)	3257 (62.1)
History of angina	1559 (26.9)	1424 (27.2)
History of claudication	390 (6.7)	354 (6.8)
History of myocardial infarction	776 (13.4)	708 (13.5)
History of stroke or TIA	649 (11.2)	586 (11.2)
History of vascular disease*	2565 (44.2)	2336 (44.5)

*Any of stable angina, intermittent claudication, stroke, transient ischemic attack, myocardial infarction, peripheral artery disease surgery, or amputation for vascular disease more than 6 months before study entry.

Table 2. Genomic loci with a genome wide significant p-value $\leq 5.0e-08$ in the association with baseline LDL cholesterol levels

Chr.	Gene	Nr of SNPs	Top SNP	Variant	MAF	Beta	SE	p-value	Ref*
19	APOE/APOC1	17	rs445925	G>A	0.11	-0.33	0.03	2.8e-30	7;11-14;18;19
19	LDLR	5	rs6511720	G>T	0.13	-0.19	0.02	5.2e-15	7;11;13;14;19
5	HMGCR	5	rs258494	G>C	0.38	0.10	0.02	1.3e-09	7;11;13;14;19
11	FADS2/FEN1	14	rs174541	C>T	0.38	-0.10	0.02	1.1e-08	7;11;13;19
1	PSRC1/CELSR5	1	rs602633	G>T	0.23	-0.11	0.02	5.0e-08	7;11-14;16-19

Abbreviations: SNP, Single Nucleotide Polymorphism; Chr, Chromosome; MAF, minor allele frequency; SE, standard error. * A list of references in which the same loci were found.

Table 3. Replication of the 5 significant loci in the WOSCOPS trial and CARE study in association with baseline LDL cholesterol levels

SNP	Gene	Chr.	WOSCOPS (n=431)			CARE (n=751)		
			beta	se	p-value	beta	se	p-value
rs445925	APOE/ APOC1	19	0.07	0.05	0.164	-0.10	0.04	0.006
rs6511720	LDLR	19	-0.03	0.05	0.657	-0.03	0.03	0.411
rs258494* ¹	HMGCR	5	0.06	0.03	0.044	0.03	0.02	0.147
rs174541* ²	FADS2/FEN1	11	-0.04	0.03	0.264	-0.03	0.02	0.134
rs602633* ³	PSRC1/CELSR5	1	-0.09	0.04	0.026	-0.05	0.02	0.035

* A proxy for this SNP was used in both replication cohorts, for ¹ the proxy SNP was rs7715806 with a r^2 of 0.93, for ² the proxy SNP was rs174545 with a r^2 of 0.90, and for ³ the proxy SNP was rs660240 with a r^2 of 0.88.

Abbreviations: SNP, Single Nucleotide Polymorphism; Chr, Chromosome.

association study with baseline LDL cholesterol levels within the PROSPER/PHASE study are depicted in a Manhattan plot. Forty-two SNPs in five genomic loci, *APOE/APOC1*, *LDLR*, *FADS2/FEN1*, *HMGCR*, and *PSRC1/CELSR5*, reached the genome-wide significant p-value of 5.0×10^{-8} . In table 2 a summary of the five genomic loci and their corresponding SNPs is given. The top SNP (rs445925, Chr. 19) with a p-value of $p=2.8 \times 10^{-30}$ is located within the *APOC1* gene and near the *APOE* gene. Sixteen other SNPs in the same genomic region were also found to be associated with LDL cholesterol levels. The second top SNP (rs6511720, Chr. 19) with a p-value of $p=5.22 \times 10^{-15}$ is located within the *LDLR* gene. The three other genomic regions included the *HMGCR* (Chr.5), *FADS2/FEN1* (Chr. 11), *PSRC1/CELSR5* (Chr. 1) genes. All 5 genomic loci were previously found in association with LDL cholesterol levels and no novel loci were identified.

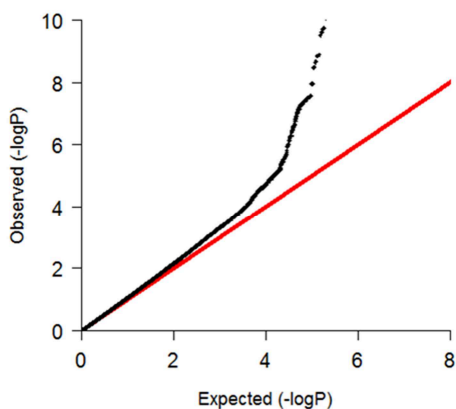


Figure 2. QQ-plot for the GWAS on baseline LDL cholesterol in the PROSPER/PHASE study.

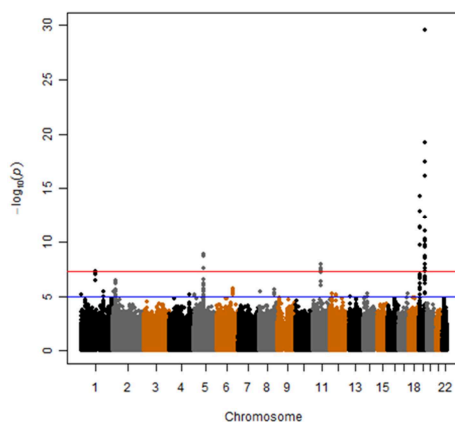


Figure 3. Manhattan plot for the GWAS on baseline LDL cholesterol in the PROSPER/PHASE study.

We replicated the positive associations with genome-wide significant p-values in two independent cohorts, the WOSCOPS study and the CARE trial (table 3). Of our five genomic loci that were significantly associated with baseline LDL cholesterol levels we selected the top SNP for replication in both replication cohorts. If the SNP was not genotyped in their GWAS analysis, we chose a proxy in high linkage disequilibrium ($r^2 > 0.5$) for that SNP. These SNPs were associated with baseline LDL levels before randomisation to statin treatment in both studies. Three out of the five loci (*APOE/APOC1*; *HMGCR*; *PSRC1/CELSR5*) replicated in one or two replication cohorts ($p < 0.05$). The two other loci (*LDLR* and *FADS2/FEN1*) showed similar trends as shown in the discovery cohort, although they did not reach statistical significance (table 3).

Discussion

With this first proof-of-principle study we show that the PROSPER/PHASE GWAS can confirm previously found genetic associations with LDL cholesterol levels. This proof-of-principle study indicates that the PROSPER/PHASE study is likely to be capable of detecting genomic regions responsible for the variation in various other quantitative traits. With almost 6000 samples in the PROSPER/PHASE study and access to various replication studies, the PROSPER/PHASE study can provide a good testing frame to identify the genetic variation responsible for the variation in LDL cholesterol lowering in response to statin treatment.

The main locus responsible for the person-to-person variation in LDL cholesterol levels is the chromosome 19 locus, which contains the *APOE*, *APOC1*, and *LDLR* genes. Other important loci included the *HMGCR* locus on chromosome 5, *FADS2/FEN1* locus on chromosome 11, and the *PSRC1/CELSR5* locus on chromosome 1. The five genomic loci that were associated with variation in LDL cholesterol levels in the PHASE GWAS study were all genomic regions that were previously reported with LDL cholesterol variation^{7;11-19}. Three out of the five loci were replicated in the WOSCOPS study and the CARE trial. The *LDLR* and *FADS2/FEN1* loci were not replicated, however these loci were repeatedly found to be associated with LDL cholesterol levels in various other studies with large number of participants^{7;11-14;16;19}. Moreover, both the WOSCOPS and CARE studies had genotype data available in a small number of subjects. Therefore, the lack of replication of these loci in WOSCOPS and CARE was most likely due to lack of statistical power. Finally, since we used in the replication studies a proxy SNP for some of the topSNPs, this may have diluted the effect.

Conclusions

With this proof-of-principle study we show that the PROSPER/PHASE study can be used to investigate genetic associations in a similar way to population based studies. Moreover, we can also assume from these results that the PROSPER/PHASE study is likely to have sufficient power to detect genome-wide significant hits with large effects for other quantitative traits. The next step of the PROSPER/PHASE study is to identify the genetic variation responsible for the variation in LDL cholesterol lowering in response to statin treatment.

Acknowledgements and Funding

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-223004. For a part of the genotyping we received funding from the Netherlands Consortium of Healthy Aging (NGI: 05060810). This work was performed as part of an ongoing collaboration of the PROSPER study group in the universities of Leiden, Glasgow and Cork. Prof. Dr. J.W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (2001 D 032).

References

- (1) McGovern PG, Pankow JS, Shahar E, Doliszny KM, Folsom AR, Blackburn H, et al. Recent trends in acute coronary heart disease--mortality, morbidity, medical care, and risk factors. The Minnesota Heart Survey Investigators. *N Engl J Med* 1996 Apr 4;334(14):884-90.
- (2) Kalantzi KJ, Milionis HJ, Mikhailidis DP, Goudevenos JA. Lipid lowering therapy in the elderly: is there a benefit? *Curr Pharm Des* 2006;12(30):3945-60.
- (3) Kreisberg RA, Oberman A. Clinical review 141: lipids and atherosclerosis: lessons learned from randomized controlled trials of lipid lowering and other relevant studies. *J Clin Endocrinol Metab* 2002 Feb;87(2):423-37.
- (4) Johnson JA, Cavallari LH. Cardiovascular pharmacogenomics. *Exp Physiol* 2005 May;90(3):283-9.
- (5) Goldstein DB, Tate SK, Sisodiya SM. Pharmacogenetics goes genomic. *Nat Rev Genet* 2003 Dec;4(12):937-47.
- (6) Shepherd J, Blauw GJ, Murphy MB, Bollen EL, Buckley BM, Cobbe SM, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* 2002 Nov 23;360(9346):1623-30.
- (7) Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010 Aug 5;466(7307):707-13.
- (8) Shepherd J, Blauw GJ, Murphy MB, Cobbe SM, Bollen EL, Buckley BM, et al. The design of a prospective study of Pravastatin in the Elderly at Risk (PROSPER). PROSPER Study Group. PROSpective Study of Pravastatin in the Elderly at Risk. *Am J Cardiol* 1999 Nov 15;84(10):1192-7.
- (9) Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 1995 Nov 16;333(20):1301-7.
- (10) Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med* 1996 Oct 3;335(14):1001-9.
- (11) Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009 Jan;41(1):47-55.
- (12) Barber MJ, Mangravite LM, Hyde CL, Chasman DI, Smith JD, McCarty CA, et al. Genome-wide association of lipid-lowering response to statins in combined study populations. *PLoS One* 2010;5(3):e9763.
- (13) Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008 Feb;40(2):189-97.

- (14) Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009 Jan;41(1):56-65.
- (15) Ma L, Yang J, Runesha HB, Tanaka T, Ferrucci L, Bandinelli S, et al. Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the Framingham heart study data. *BMC Med Genet* 2010;11:55.
- (16) Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, et al. LDL cholesterol concentrations: a genome-wide association study. *Lancet* 2008 Feb 9;371(9611):483-91.
- (17) Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* 2008 Jan;82(1):139-49.
- (18) Waterworth DM, Ricketts SL, Song K, Chen L, Zhao JH, Ripatti S, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2010 Nov;30(11):2264-76.
- (19) Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008 Feb;40(2):161-9.