POPULATION-BASED IMPACT OF SMOKING, DRINKING, AND GENETIC FACTORS ON HDL-CHOLESTEROL LEVELS IN J-MICC STUDY PARTICIPANTS

by:

Yora Nindita
(4513800183)

Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTORATE OF PHILOSOPHY

in the field of Medical Science

DEPARTMENT OF INTERNATIONAL ISLAND AND COMMUNITY MEDICINE
GRADUATE SCHOOL OF MEDICAL AND DENTAL SCIENCES
KAGOSHIMA UNIVERSITY
KAGOSHIMA, JAPAN
2021

DECLARATION

This dissertation is submitted for the degree of Doctor of Philosophy in Medical Sciences at Kagoshima University, Japan. The research described herein was conducted under the supervision of Professor Toshiro Takezaki in the Department of International Island and Community Medicine, Graduate School of Medical and Dental Sciences, Kagoshima University, between October 2013 and September 2017.

I hereby declare that this dissertation is to the best of my knowledge original and based on my original article, "Nindita Y, Nakatochi M, Ibusuki R, Shimoshikiryo I, Nishimoto D, Shimatani K, Takezaki T, Ikezaki H, Murata M, Hara M, Nishida Y, Tamura T, Hishida A, Nagayoshi M, Okada R, Matsuo K, Ito H, Mikami H, Nakamura Y, Otani T, Suzuki S, Koyama T, Ozaki E, Kuriki K, Takashima N, Miyagawa N, Arisawa K, Katsuura-Kamao S, Momozawa Y, Kubo M, Takeuchi K, and Wakai K, for the Japan Multi-Institutional Collaborative Cohort Study Group (2021). Population-Based Impact of Smoking, Drinking, and Genetic Factors on HDL-Cholesterol Levels in J-MICC Study Participants. *J Epidemiol* https://doi.org/10.2188/jea.JE20210142 (in press)" in which Introduction, Material and Methods, Results, and Discussion chapters were modified with additional information, and a Literature Review chapter was added. Neither this, nor any substantially similar dissertation has been or being submitted for any other degree, diploma, or other qualification at any other university.

Kagoshima, December 2021

Yora Nindita

ABSTRACT

Background: Environmental and genetic factors are suggested to exhibit factor-based association with HDL-cholesterol (HDL-C) levels. However, the population-based effects of environmental and genetic factors have not been compared clearly. We conducted a cross-sectional study using data from the Japan Multi-Institutional Collaborative Cohort (J-MICC) Study to evaluate the population-based impact of smoking, drinking, and genetic factors on low HDL-C.

Method: Data from 11,498 men and women aged 35-69 years were collected for a genome-wide association study (GWAS). Sixty-five HDL-C-related SNPs with genome-wide significance ($P < 5 \times 10^{-8}$) were selected from the GWAS catalog, and seven representative SNPs were defined, and the population-based impact was estimated using both 1) population attributable fraction (PAF) on low HDL-C; and 2) the similar exposure levels between smoking and drinking, and seven representative SNPs on whole levels of HDL-C.

Results: We found that smoking, drinking, daily activity, habitual exercise, egg intake, BMI, age, sex and the SNPs *CETP* rs3764261, *APOA5* rs662799, *LIPC* rs1800588, *LPL* rs328, *ABCA1* rs2575876, *LIPG* rs3786247, and *APOE* rs429358

were associated with HDL-C levels. The gene-environmental interactions on smoking and drinking were not statistically significant. The PAF for low HDL-C was the highest in men (63.2%) and in rs3764261 (31.5%) of the genetic factors, and the PAFs of smoking and drinking were 23.1% and 41.8%, respectively. As another population-based impact, the association coefficient (8.012) of the combination of the 7 SNPs (21.3% prevalence) with the HDL-C levels was 1.07 times greater than that for drinking (7.498, 22.0%), and 1.46 times greater than that for smoking (-5.447, 19.7%).

Conclusion: The present study showed that the PAF, as a population-based impact, of genomic factor *CETP* rs3764261 for low HDL-C was higher than that of smoking and lower than that of drinking.

Keywords: HDL-cholesterol, drinking, smoking, single nucleotide polymorphism, gene-environmental interaction

DEDICATION

I dedicate this work to

my parents Hardjono Hardjowitjitro and Sri Rahayu Martaningsih,

my brother Hita Manggala Nusa and his wife Lady Ramona,

my niece Odelia Anasya Prameswari and nephew Orvinoval Putra Manggala,

my best friends Nadia Hardini, Pulong Wijang Pralampita,

Muflihatul Muniroh, Martha Ardiaria, and Arwinda Nugraheni

in memoriam of Yurika Maryanti,

my colleagues Noor Wijayahadi, M. Ali Sobirin,

Nani Maharani, Endang Mahati, and Astika Widy Utomo,

and all of my families and friends in Indonesia, Japan,

and other parts of the world.

I am so thankful for all the love and support you showered me all of this time.

ACKNOWLEDGMENTS

First of all thank God Almighty; the most merciful, beneficent and compassionate; for blessing these study to be completed.

I would like to express my sincere gratitude to Toshiro Takezaki, Professor of Department of International Island and Community Medicine, for his invaluable guidance and insights throughout my doctoral study as well as in the writing and the completion of this paper.

Thank you to Dr. Nobuyuki Hamajima and Dr. Hideo Tanaka for initiating and organizing the J-MICC Study as former principal investigators.

Thank you to all of the co-authors Masahiro Nakatochi, Rie Ibusuki, Ippei Shimoshikiryo, Daisaku Nishimoto, Keiichi Shimatani, Toshiro Takezaki, Hiroaki Ikezaki, Masayuki Murata, Megumi Hara, Yuichiro Nishida, Takashi Tamura, Asahi Hishida, Mako Nagayoshi, Rieko Okada, Keitaro Matsuo, Hidemi Ito, Haruo Mikami, Yohko Nakamura, Takahiro Otani, Sadao Suzuki, Teruhide Koyama, Etsuko Ozaki, Kiyonori Kuriki, Naoyuki Takashima, Naoko Miyagawa, Kokichi Arisawa, Sakurako Katsuura-Kamao, Yukihide Momozawa Michiaki Kubo, Kenji

Takeuchi, and Kenji Wakai, for the Japan Multi-Institutional Collaborative Cohort Study Group.

Thank you to the local governments and staffs from Chiba, Shizuoka, Aichi, Shiga, Kyoto, Tokushima, Fukuoka, Saga, Nagasaki, Kagoshima, and Okinawa for collaborating in the data collection.

Sincere regards to former department member, Eva Mariane Mantjoro and Tara Sefanya Kairupan for kind help and cooperation during field and laboratory work.

At last, I want to thank all of the professors, lecturers, and administrative staffs of Graduate School of Medical and Dental Science, Kagoshima University for providing me so much knowledge along my doctoral study. To my almamater Diponegoro University, Gadjah Mada University and to the Ministry of Research and Higher Education Republic of Indonesia, thank you for giving me the chance to pursue higher education abroad.

This study was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas of Cancer (no. 17015018) and Innovative Areas (no. 221S0001) and by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant (No. 16H06277 [CoBiA]) from the Japanese Ministry of Education, Culture, Sports,

Science and Technology. This study was also supported in part by funds for the BioBank Japan Project received from the Japan Agency for Medical Research and Development since April 2015, and from the Ministry of Education, Culture, Sports, Science and Technology from April 2003 to March 2015.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA : Alterative Homo-Genotype

ABCA1 : ATP-binding cassette transporter A1

ADH : Alcohol Dehydrogrnase

Apo : Apolipoprotein

BMI : Body Mass Index

CETP : Cholesteryl Ester Transfer Protein

CIs : Confidence Intervals

Coeff. : Coefficient

CVD : Cardiovascular Disease

DNA : Deoxyribonucleic Acid

GA : Hetero-Genotype

GALNT2 : GalNac Transferase

GG : Homo-Genotype

GWAS : Genome-Wide Association Study

H_2O_2 : Hydrogen Peroxide

HDL-C : HDL-cholesterol

IDL : Intermediate-density lipoprotein

J-MICC : Japan Multi-Institutional Collaborative Cohort

LCAT : Lecithin Cholesterol Acyltransferase

LDL : Low-Density Lipoprotein

LIPC : Hepatic Lipase

NMR : Nuclear Magnetic Resonance

MAF : Minor Allele Frequency

MEOS : Microsomal Ethanol Oxidizing System

METs : Metabolic Equivalents

OH : Hydroxil

OR : Odd Ratio

PAF : Population Attributable Fraction

PCA : Principal Component Analysis

RA : Referent And Alterative Allele Hetero-Genotype

RCT : Reverse Cholesterol Transport

REF : Referent Allele

RR : Referent Allele Homo-Genotype

SCARB1 : Scavenger Receptor Class B Type I

SNPs : Single Nucleotide Polymorphism

VLDL : Very-Low-Density Lipoprotein

CHAPTER 1: INTRODUCTION

1.1 Background

Low serum levels of HDL-cholesterol (HDL-C) are associated with an increased risk of cardiovascular disease (CVD).[1,2] As there are no clinically available drugs that can enhance HDL-C levels, along with genetic factors, environmental factors too play an important role in the alleviation of CVD risk. Smoking, alcohol intake, physical activity, body mass index (BMI), and diet intake have been confirmed to be environmental factors that affect HDL-C levels.[2-6]

The effects of genetic factors, such as single nucleotide polymorphism (SNPs) in various enzymes-encoding genes, on HDL-C levels have also been reported.[7] Although the regulation of HDL-C metabolism is a complex process, enzymes in the reverse cholesterol transport (RCT) system, such as ATP-binding cassette transporter A1 (ABCA1), lecithin cholesterol acyltransferase (LCAT), cholesteryl ester transfer protein (CETP), hepatic lipase (LIPC), APOA1/C3/A4/A5, scavenger receptor class B type I (SCARB1), and LPL play a major role in it.[2] Multiple SNPs have been reported to be associated with HDL-C levels, and among the genes harboring such SNPs, the genetic variants of *CETP* have been observed to exert a

greater influence on HDL-C levels.[8-11] Furthermore, besides association with SNPs in RCT-related genes, the association with several other SNPs, such as those in genes encoding endothelial lipase (LIPG) and APOE, which are related to lipoprotein dynamism, has been reported.[10,12]

The majority of studies on environmental and genetic factors that affect HDL-C levels focus on factor-based association with respect to individual risk and susceptibility, and the population-based impact of environmental and genetic factors on HDL-C levels has not been discussed clearly. The population-based impact of a factor is an important aspect for public health. The population-based impact of various environmental factors on HDL-C levels can be estimated based on the impact of the association and prevalence of each factor. However, the population-based impact of genetic factors is difficult to estimate, because several SNPs are detected in each enzyme-encoding gene; the impact of the association of each SNP with HDL-C levels will differ, and the prevalence of the allele containing each SNPs will differ as well. Therefore, studies that investigate the combined effect of HDL-C-related SNPs limit their assessment to certain representative SNPs.[9] Furthermore, gene-environment interaction may influence HDL-C levels as well.[13,14]

Among environmental factors, smoking and drinking habits significantly affect the reduction or increase in HDL-C levels, respectively.[2,9,15] These factors are suitable candidates for the estimation of the population-based impact of environmental factors on HDL-C levels, taking into account the interaction with genetic factors as well. In such cases, GWAS are suitable for evaluating the overall scenario. GWAS on the effects of HDL-C-related SNPs on ethnic populations, including the Japanese population, have been performed earlier, and all HDL-C-related SNPs have been listed in the catalog.[16,17]

1.2 Aim

To investigate the population-based impact of smoking, drinking, and genetic factors on low HDL-C, we conducted a relatively large-sized cross-sectional study using data on environmental factors and GWAS from the Japan Multi-Institutional Collaborative Cohort (J-MICC) Study.

CHAPTER 2: LITERATURE REVIEW

2.1 HDL-C Overview

HDL is beneficial cholesterol for human body due to it transports cholesterol and prevents the deposition of cholesterol in the arteries (atherosclerosis). HDL has a less fat content and has a high density (1.063–1.21 g/ml) with a size 8-10 nm. The main protein that makes up HDL is Apo-A (apolipoprotein). Normal HDL cholesterol should be higher than 40 mg/dL for men, or above 50 mg/dL for women.[18] Lower HDL level is associated with increased risk of CVD, such as myocardium infarction, stroke, stenosis, and atherosclerotic. Thus, level of HDL might be used for calculation of risk factor of CVD.[19,20]

2.1.1 HDL Subclasses

Several factors were affecting the heterogeneity of HDL resulting in different subclasses (Table 1). The subclasses of HDL are grouped according to their density, size, shape, shape and charge, size and charge, and protein composition. Each grouping is done with a different separation method as shown in table below.[2,21]

Table 1. Major HDL subclasses according to different separation technique[2]

Density (ultracentrifugation)
HDL2 (1.063–1.125 g/mL)

HDL3 (1.125–1.21 g/mL)

Size (Gradient Gel Electrophoresis)

HDL2b (9.7–12.0 nm)

HDL2a (8.8–9.7 nm)

HDL3a (8.2–8.8 nm)

HDL3b (7.8–8.2 nm)

HDL3c (7.2–7.8 nm)

Size (Nuclear Magnetic Resonance)

Large HDL (8.8–13.0 nm)

Medium HDL (8.2–8.8 nm)

Small HDL (7.3–8.2 nm)

Shape and charge (agarose gel)

α-HDL (spherical)

Preβ-HDL (discoidal)

Charge and size (2D electrophoresis)

Preβ-HDL (preβ1 and preβ2)

α-HDL (α1,α2,α3 andα4)

Preα-HDL (preα1, preα2, preα3)

Protein composition (electro immunodiffusion)

LpA-I

LpA-I:A-II

2.1.2 HDL Structure and Formation

The most common formation of HDL in plasma is spherical. In spherical formation, the core consists of triglycerides and cholesteryl esters. Without cholesteryl esters, HDL has a discoidal shape.[22] The surface component of HDL are surrounded by phospholipids and apoproteins. Phospholipids, mainly

phosphatidylcholine, is the major lipid in HDL (35-50%). Apoprotein A-1 (apoA-1) is the major apoprotein in HDL (70%), and apoA-II (20%).[23]

2.1.3 HDL Biogenesis and Catabolism

Lipid-free apoA-1 is produced in the liver (70%) and small intestine (30%).[24,25] In peripheral cells, the cholesterol and phospholipids will be transfer into lipid-free apoA-1 by ABCA1, resulting in discoidal shape of HDL. Esterification of cholesterol by enzyme lecithin cholesterol acyltransferase (LCAT) transforms the HDL into a spherical shape and produced HDL3 and HDL2. In addition to ABCA1, ABCG1 also helps cholesterol efflux for esterification. The cholesterol can be directly transferred to the liver through SR-B1/SCARB1 or indirectly transferred to very-low-density lipoprotein (VLDL), intermediate-density lipoprotein (IDL), and low-density lipoprotein (LDL) by CETP.

This indirect pathway product will be transferred to the liver through LDLR.[26,27] The catabolism of HDL is through uptake of HDL2 to the liver by HDLR. The hepatic, endothelial lipase enzyme and plasma phospholipid transfer protein transform into lipid-poor HDL particles and the subsequent cycle is

repeated.[28]

2.1.4 Laboratory screening and confirmatory tests

The laboratory test is useful to prevent CVD in healthy populations, early detection of risky individuals, and slow down the progression of hypercholesterolemia. Cholesterol screening should be carried out every 1-2 years at the age of 45-65 years for men and 55-65 years for women. For patients >65 years of age, cholesterol screening should be performed annually. Meanwhile, other patients are indicated to conduct the laboratory test with familial history of premature death due to myocardial infarction at age < 55 years for men and <65 years for women, familial history of hypercholesterolemia, including xanthomas / Xanthomata, patients with comorbid such as obesity, diabetes, HIV, hypertension that will be beneficial with statin used, and a person with sedentary life for risk stratification.[29,30]

There were several techniques to determine HDL-C in serum, including ultracentrifugation, electrophoresis, gel permeation chromatography, proteomic, lipodomic, HPLC, capillary isotachophoresis, nuclear magnetic resonance (NMR), enzyme assay, cholesterol efflux, precipitation-based methods, and direct

measuring methods. The gold standard for lipoprotein separation is density gradient ultracentrifugation. Meanwhile, the gold standard HDL cholesterol removing capacity is cholesterol efflux method. The most widely used in clinical setting of HDL-C measurement are precipitation-based methods and direct measuring methods.[31,32]

2.2 Smoking

Tobacco is produced from the *Nicotiana tabacum, Nicotiana ristica,* and other plants which contain nicotine and tar with or without additives. Cigarette smoke brings harm from its tobacco content and also from the combustion products. About 60% of cigarette smoke consists of gases and vapors, including carbon monoxide, hydro cyanide, nitric acid, fluorocarbon nitrogen dioxide, acetone, and ammonia.[33,34]

Smoking behaviour is an activity of sucking tobacco smoke from a burned cigarette into the body and exhaling it back out. A smoker is someone who inhales cigarette smoke either directly through cigarettes or not. An active smoker is someone who consumes cigarettes regularly, even if it is only one cigarette, or a person who smokes cigarettes even if not routinely or just trial and error. While

passive smokers are people who are not smokers but participate in inhaling other people's cigarette smoke.[35,36]

2.2.1 Relation of Smoking and HDL

Smoking is widely known to cause major health problem especially CVD. Moreover, cigarette smoking is associated with reduced HDL cholesterol levels. Smoking generates higher oxidative stress in the body and this oxidative may alter the HDL function and lose its atheroprotective properties in smokers. Smoking may lead to the reduction of LCAT, CETP, and hepatic lipase activity.[6,37]

2.3 Drinking

A group of chemical compounds that have one hydroxyl (–OH) group are alcohols. Methanol, ethanol, and isopropanol are types of alcohol that are often found and are widely used as solvents. Alcoholic beverages are drinks containing alcohol and ethanol (C2H5OH) derived from the fermentation process of various plant raw materials such as seeds, fruits, or sap containing carbohydrates which are the result of the fermentation process or distillation (distillation). Ethanol (C2H5OH) is an alcoholic solution found in alcoholic beverages in general.[38]

Alcoholism is a term for people who consume and are addicted to alcohol. Alcohol addiction is a complex disorder and often seen from a biopsychosocial perspective. An alcoholic is someone who is addicted to alcohol, which is characterized by excessive drinking and subsequent with mental and social disorders. The tendency of alcoholism behaviour is not similar in every individual. Not everyone who consumes alcohol will become an alcoholic.[39,40]

2.3.1 Alcohol Metabolism

Alcohol is metabolized into the form of acetaldehyde compounds in the body in 2 ways[41], alcohol dehydrogrnase (ADH) and microsomal ethanol oxidizing system (MEOS).

The main pathway of alcohol metabolism involves the enzyme ADH, a group of cytosolic enzymes that catalyze the conversion of alcohol to acetaldehyde. This enzyme is abundant in the liver but is also found in the brain and stomach in small amounts. During the conversion of ethanol by ADH to acetaldehyde, hydrogen ions are transferred from the ethanol to the cofactor nicotinamide adenine dinucleotide (NAD+) for the formation of NADH. There are excess reducing equivalents in the liver resulting from alcohol oxidation. Excess NADH production can contribute to

metabolic disorders that occur in chronic alcoholism. Meanwhile, acute alcoholism can cause hypoglycaemia or cause lactic acidosis.[41,42]

Mixed function oxidizing system, also known as MEOS, in ethanol metabolism uses NADPH as a cofactor consisting of cytochrome P450 (CYPs) such as CYP2E1, CYP1A2, and CYP3A4. Induction of MEOS activity due to chronic alcohol consumption. In addition, other drugs carried out by cytochrome P450 in the MEOS system will be affected, as well as the formation of toxic products and reactions of cytochrome P450 such as toxins, free radicals, and hydrogen peroxide (H2O2).[41,43]

2.3.2 Alcohol-related disease

Alcoholic liquor is a factor that causes about 60 types of disease and is a component factor of 200 other diseases. Liver dysfunction, such as alcoholic liver disease, is a type of disease caused by alcohol consumption. Excessive consumption of alcohol for a long period can cause increase blood pressure and persist into hypertension, cardiovascular damage, stroke, breast cancer, liver damage, digestive tract cancer, and other digestive disorders. Alcohol can also cause impotence and reduced fertility, sleep difficulty, brain damage followed by changes in personality and mood, memory and concentration disorders.[41,44]

2.3.3 Alcohol and HDL-C

Alcohol is known to affect on the metabolism of HDL-C, LDL-C, and triglycerides. Alcoholics consumption may lead to myocardial infarction, decreased levels of LDL-C and fibrinogen, and increased levels of HDL-C. The mechanism of alcohol increase HDL-C is might be through an increase in transport rate of apoA-1 and apoA-II. Moreover, alcohol also might contribute to cholesterol esterification by stimulating the efflux of cholesterol from peripheral cell.[45,46]

2.4 Genome-Wide Association Study (GWAS)

The Genome-Wide Association Study (GWAS) is an approach that involves the complete set of deoxyribonucleic acid (DNA)/genomes from several of individuals who are affected by a disease and not affected by the disease to find genetic variations associated with the disease. GWAS is a powerful tool to investigate genetic variants that contribute to complex human traits. GWAS is a more cost-efficient way to survey common genetic variation compared to candidate gene approach. [47,48]

To carry out genetic testing on diseases for which the genes are well established, scientists collect blood samples from a group of individuals with the disease and

analyze their DNA for SNP patterns. SNP are the most common variation in general population due to it is contain two alleles, resulting in two possibilities of base-pair. After the SNP pattern is collected, researchers compared these patterns with patterns obtained by analyzing DNA from a group of individuals who did not develop the disease. Thus, it can be shown which patterns are most likely to be associated with disease-causing genes.[49,50]

For diseases in which the causative gene is unknown, GWAS study can determine candidate genes and the associated SNPs. A key aspect of GWAS is the association of SNPs with inherited phenotypes. SNP studies can accelerate the identification of disease genes by enabling researchers to look for associations between disease and SNP differences in individuals within a population.[49,50]

In this study, large-scale genotyping platforms are used to assay hundreds of thousands of SNPs simultaneously.[51,52] GWAS have identified numerous common variants associated with HDL-C, LDL-C, and TG.[53] Recently, GWAS have revealed common genetic factors that are associated with the HDL-C concentration in blood and demonstrated strong evidence for the ABCA1, apolipoprotein, CETP, GalNac transferase (GALNT2), LIPG, and LPL gene polymorphisms.[16,54-58]

CHAPTER 3: METHODS

3.1 Study Design

We used a cohort study design to investigate the interaction effect between all SNPs and HDL-cholesterol levels; and between smoking and drinking habits and selected SNPs for HDL-cholesterol levels with the genetic background with GWAS approach among the general population in Japan.

This research is a part of the Japan Multi-institutional Collaborative Cohort (J-MICC) study. The J-MICC study was first conducted in 2005, supported by a research grant for Scientific Research on Special Priority Areas of Cancer from the Japanese Ministry of Education, Culture, Sports, Science, and Technology. The primary goal of the J-MICC study is to confirm and detect gene-environment interactions in the development of lifestyle-related diseases, mainly of cancer, through the cohort analyses. [59,60]

The ethics committees of Nagoya University Graduate School of Medicine, Aichi Cancer Center, Kagoshima University Graduate School of Medical and

Dental Sciences, and other all participating institutes and universities approved the protocol.

3.2 Study Region

The subjects (n=14,555) of the GWAS selected from among the J-MICC Study participants were aged between 35-69 years and belonged to 11 prefectures of Japan (Chiba, Shizuoka, Aichi, Shiga, Kyoto, Tokushima, Fukuoka, Saga, Nagasaki, Kagoshima, Okinawa), participants were selected by ten research institutes and universities.

3.3 Population and Sample

The participants were recruited from health-checkup examinees by local government, private companies and health-checkup center; responders by posting method to regional residents; and first-visit outpatients at cancer center. All subjects (n=14,555) from 11 prefectures selected using several criteria. The present exclusion criteria were the lack of information on HDL-C levels (all participants [n=2,296] from Cancer Institute of Chiba study region and Aichi Cancer Centre, and participants [n=187] from others) and, smoking (n=180) and drinking (n=24)

habits; and from cases of withdrawal (n=21) were excluded as well. Data from certain subjects qualified for multiple exclusion criteria. The final number of eligible subjects was 11,498 (the dataset used in the present study was decided upon on March 12, 2020, version 20200312).

3.4 Data Collection

3.4.1 Questionnaire

A standardized structured questionnaire was used in the J-MICC Study to collect information regarding lifestyle factors and medical history of the subjects.[61] The questionnaire was evaluated by trained staff to ensure completeness and consistency.

3.4.2 HDL-C assessment

Venous blood samples were collected from the subjects in sitting position during a period of fasting. The mean duration of fasting was 9.8 h. The blood samples were separated into serum, plasma, and buffy coat fractions, and stored directly at -80 °C on the day of sampling. The serum HDL-C levels were measured at the respective institutes for health checkup or medical examination in each study region.[62]

3.4.3 Quality of samples and SNPs on genotyping

DNA was extracted from the buffy coat fractions using a BioRobot M48 Workstation (Qiagen Group, Tokyo, Japan) at Nagoya University, using samples from all regions except Fukuoka and KOPS (Kyushu and Okinawa Population Study); DNA was extracted from the samples from these two regions at Kyushu University using an automatic nucleic acid isolation system (NA-3000, Kurabo, Co., Ltd, Osaka, Japan). Next, the DNA samples were genotyped at the RIKEN Center for Integrative Medicine using a HumanOmniExpressExome-8 v1.2 BeadChip array (Illumina Inc., San Diego, CA, USA). The number of low-quality DNA samples was 463, which were excluded from the analysis. The subjects for whom sex information in the questionnaire was inconsistent with that revealed by the genotyping results were excluded. Furthermore, the identity-by-descent method implemented in the PLINK 1.9 software[63] was used to identify close relationship pairs (pi-hat > 0.1875) and the sample from each pair was excluded. The subjects (n=34) with non-Japanese estimated ancestries[64] were also excluded by principal component analysis (PCA)[65] using a 1000 Genomes reference panel (phase 3).[66]

SNPs with a genotype call rate <0.98, a Hardy-Weinberg equilibrium exact test P-value < 1×10^{-6}, and a low minor allele frequency (MAF) < 0.01, or a

departure from the allele frequency computed from the 1000 Genomes Phase 3 EAS (East Asian) samples; and non-autosomal SNPs were excluded. Such quality control filtering resulted in 14,091 individuals and 570,162 SNPs.

3.4.4 Genotype imputation and post-imputation quality control

The imputation of genotypes in autosomal chromosomes was performed using SHAPEIT2[67] and Minimac3[68] software with the 1000 Genomes reference panel (phase 3).[66] The imputation procedure displayed 47,109,431 SNPs from 570,162 SNPs.

The SNPs with imputation quality $r^2 < 0.3$ were excluded in the post-imputation quality control step. The number of eligible SNPs was 12,617,547.

3.4.5 Selection of HDL-C related SNPs

On August 27, 2019, HDL-C-related SNPs were systematically selected from the GWAS catalog (https://www.ebi.ac.uk/gwas/) (the database of published GWAS), which included 499 SNPs from all ethnic population.[16, 17] Next, 65 SNPs among these were selected for the present study, which had *P*-values of genome-wide significance ($P < 5 \times 10^{-8}$) in the present analysis (Table A.1). The Q-Q plot showed the apparently different distribution of the present observed $\log_{10}$ (*P*-value)

of the 65 SNPs against the expected $\log_{10}$ (*P*-value) (Figure 1). Although the association for rs921919 in *SCARB1* (12q24.31) indicated genome-wide significance, this was not included in the present analysis because this SNP was not previously reported to be associated with HDL-C levels and were not listed in the GWAS catalog. Other SNPs in *SCARB1* listed in the GWAS catalog were not genome-wide significant in the present analysis.

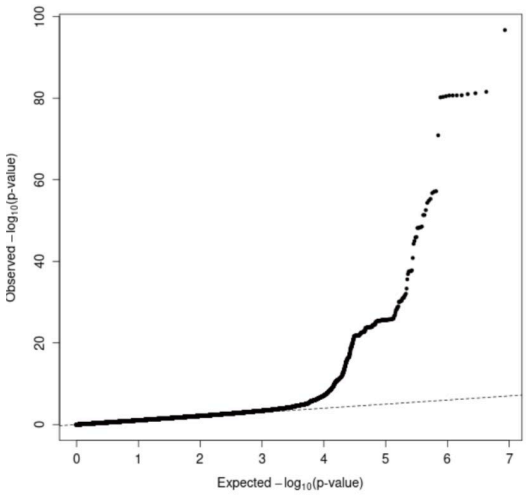

Figure 1. Q-Q plot for *P* values from original GWAS data. The vertical and horizontal axes indicate observed and expected $-\log_{10}$ (*P* value) for tests of association between SNPs and HDL-C, respectively.

3.5 Statistical analysis

The subjects were divided into two categories based on the smoking status ("never" and "former" [≥ 1 year] vs. "current" [include smokers within 1 year after quitting]), because the HDL-C levels apparently differed between subjects with the "current" and "never" statuses, and with respect to the duration after quitting. The subjects were also divided into two categories based on the drinking status (non-, former, and current moderate drinkers [<20 g/day] vs. current heavy drinkers [≥20 g/day]), as the Japanese Ministry of Health, Labor and Welfare recommends alcohol intake in moderation (at <20 g/day); the HDL-C levels apparently differed between the two categories.[68] The duration and intensity of daily activity (hard work and walking) and the frequency and intensity of habitual exercise were used to estimate the metabolic equivalents (METs). The estimation of METs・hour per day was based on the duration and intensity of exercise, with 3.0 for walking, 3.3 for light exercise, 4.0 for moderate exercise, 4.5 for heavy work, and 8.0 for heavy exercise.[69] Daily activity was classified as < 8.25 METs・h/day and ≥ 8.25 METs・h/day at the median value. Habitual exercise was classified as < 0.728 METs・h/day and ≥ 0.728 METs・h/day at the median value. Egg intake was selected as a representative HDL-C-related dietary factor.[2, 9] There were two categories for BMI with comparable number of male and female subjects in each. The association

between HDL-C levels (continuous) and non-genetic factors such as smoking and drinking habits was tested using multivariate linear regression analysis after adjusting for the following HDL-C-related factors: age (<57 vs. ≥57 years), sex, smoking, drinking, daily activity, habitual exercise, egg intake and BMI. Dummy variables of 0 and 1 were used for all independent variables. Statistical analyses for non-genetic factors were performed using Stata software (version 12; Stata Corp., College Station, TX, USA), differences with p-value < 0.05 were considered statistically significant.

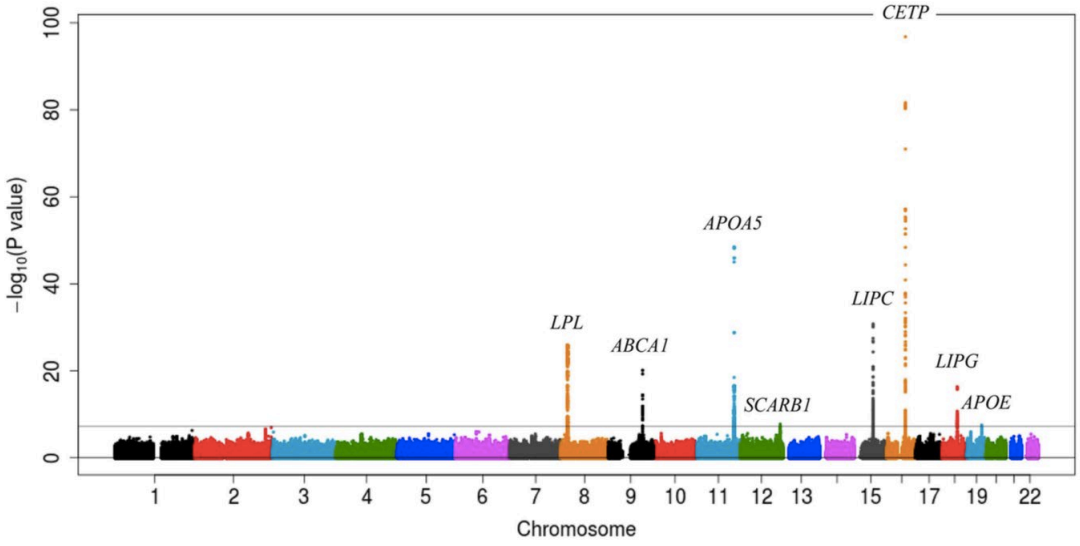

Figure 2. Manhattan plot ($-\log_{10}$ of the P value based on genomic location) of the association between the SNPs denoted in the original GWAS and the HDL-C levels shows the formation of eight peaks over the line representing $P < 5 \times 10^{-8}$ for *LPL* (8p21.3), *ABCA1* (9q31.1), *APOA5* (11q23.3), *SCARB1* (12q24.31), *LIPC* (15q21.3), *CETP* (16q13), *LIPG* (18q21.1), and *APOE* (19q13.32). The horizontal line represents the genome-wide significance level ($\alpha = 5 \times 10^{-8}$).

The selected HDL-C-related 65 SNPs were divided into seven categories based on the gene and cytoBand groups (Table A.1). The Manhattan plot for total SNPs in the present GWAS consistently showed seven peaks with genome-wide significance, with the exception of a single peak corresponding to rs921919 in *SCARB1* with genome-wide significance yet unlisted in the GWAS catalog (Figure 2). Next, the seven SNPs with the highest coefficients and lowest *P*-values from each of the seven groups were selected. The association between HDL-C levels (continuous) and genetic factors, and the interaction were tested using multivariate linear regression analysis in epacts v3.2.6 software (https://genome.sph.umich.edu/wiki/EPACTS), after adjusting for the HDL-C-related factors and first 5 principal components. Dummy variables of 0, 0.5, and 1 were used for the number of alternative alleles (0, 1 and 2) as independent variables in order to compare the impact of coefficients on non-genetic factors (dummy variables of 0 and 1), and the coefficients and 95% confidence intervals (CIs) were estimated. Differences with $\alpha = 5 \times 10^{-8}$ were considered statistically significant in the GWAS. We applied the Bonferroni correction ($P < 0.00077$) for evaluating the differences in interaction of smoking or drinking with the 65 SNPs to reduce the chances of introducing an alpha error by multiple hypothesis testing.

The population-based impact of the non-genetic and genetic factors was estimated using population attributable fraction (PAF).[31,32] First, the odd ratio (OR) for low HDL-C (<40 mg/dL) was estimated, and the PAF was calculated as;

$$PAF = P \times \frac{(OR-1)}{OR} \times 100\ (\%)$$

where P is the proportion of the exposure in subjects with low HDL-C. The reference exposure group was defined as those with the minimum risk for low HDL-C, i.e. smoking habit ("never" and "former" [≥ 1 year]), drinking habit (≥ 20 gram alcohol/day), daily activity (≥ 8.25 METs/day), habitual exercise (≥ 0.73 METs/day), egg intake (≥ 3 times/week), BMI (< 23.0), age (< 57 years) and sex (women) in the non-genetic factors; and rs3764261, rs662799, rs1800588, rs328 and rs3786247 (referent and alterative allele hetero-genotype, and alterative allele homo-genotype), and rs2575876 and rs429358 (referent allele homo-genotype) in the genetic factors. Dummy variables of 0 and 1 were used for both the non-genetic and genetic factors. When the PAF of the combined SNPs was estimated, the accumulation in 6 SNPs was categorized according to the number of the high-risk genotypes for low HDL-C by individual regardless kind of SNPs, i.e. 0-1 SNPs for reference, 2 SNPs, 3 SNPs and 4-6 SNPs. The SNP of rs1800588 was excluded from this accumulation analysis, because the OR for low HDL-C was not

statistically significant. The ORs and their 95% CIs were estimated using logistic model after adjusting for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake and BMI.

We also estimated the another population-based prevalence of the combination of SNPs while considering the impact on whole levels of HDL-C without clinical cutoff point (<40 mg/dL), we 1) weighted dummy values of the analyzed SNPs after multiplying each dummy value by the coefficient of association with HDL-C levels when the coefficient was positive; 2) changed the dummy values between referent and alternative alleles, and then weighted the dummy values after multiplying each dummy value by the absolute coefficient value when the coefficient was negative; 3) added the weighted values of the analyzed SNPs; and 4) defined cutoff values as higher 21.3% prevalence of summarized values in the subjects. The cutoff values for the two groups comprising 65 and 7 SNPs were respectively defined. The 21.3% SNP prevalence was considered concordant with the prevalence of male and female smokers (19.7%) and drinkers (22.0%) from among the subjects. Sensitivity analysis was performed for the coefficients of drinking exposure and SNP prevalence to assess the dose-response associations between the coefficients at

various levels of exposure, which were divided into two groups according to approximately 10%, 20%, and 40% prevalence.

CHAPTER 4: RESULTS

4.1 Baseline characteristics

The distribution pattern of male and female subjects in the two age groups (35-56 years and 57-69 years) was almost similar (Table 2). The prevalence of current smokers was 34.9% among male and 7.3% among female subjects (19.7% in both), and that of heavy drinkers was 42.7% in males and 5.1% in females (22.0% in both). The prevalence of low HDL-C (<40 mg/dL) was 8.9% in male and 1.8% in female subjects (5.0% in both sexes).

Drinking ($P < 0.001$), daily activity ($P < 0.001$), habitual exercise ($P < 0.001$), egg intake ($P = 0.004$), and sex ($P < 0.001$) were associated positively with the HDL-C levels, while smoking ($P < 0.001$), BMI ($P < 0.001$), and age ($P < 0.001$) were associated negatively (Table 3).

Table 2. Age-, environmental factor-, BMI-, and HDL-C level-based distribution of study subjects divided by sex

	Number (%)			
	Men		Women	
Age in years				
35-56	2,595	(50.7)	3,280	(52.4)
57-69	2,519	(49.3)	2,976	(47.6)
Total	5,114	(100)	6,256	(100)
Smoking				
Never and former (≥ 1 year) smokers	3,329	(65.1)	5,802	(92.7)
Current smokers[a]	1,785	(34.9)	454	(7.3)
Drinking				
Non-, former and moderate drinkers[b]	2,933	(57.4)	5,936	(94.9)
Heavy drinkers[c]	2,181	(42.7)	320	(5.1)
Daily activity				
<8.25 METs·h/day	3,102	(60.7)	3,484	(55.7)
≥8.25 METs·h/day	2,012	(39.3)	2,772	(44.3)
Habitual exercise				
<0.73 METs·h/day	2,447	(47.9)	3,268	(52.2)
≥0.73 METs·h/day	2,667	(52.2)	2,988	(47.8)
Egg intake				
<3 times/week	3,659	(71.6)	4,377	(70.0)
≥3 times/week	1,455	(28.5)	1,879	(30.0)
BMI				
<23	2,066	(40.4)	3,789	(60.6)
≥23	3,048	(59.6)	2,467	(39.4)
HDL-C				
<40 mg/dL	454	(8.9)	111	(1.8)
≥40 mg/dL	4,660	(91.1)	6,145	(98.2)

BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; METs, metabolic equivalents. [a] Smokers within 1 year after quitting were included. [b] <20 g alcohol/day. [c] ≥20 g alcohol/day.

Table 3. Association between HDL-C levels and environmental factors determined in multivariate regression analysis

	Coeff.[a]	95% CI	*P*-value
Smoking (current)	-5.407	-6.133 to -4.681	*<0.001*
Drinking (≥20 g alcohol/day)	7.274	6.549 to 8.000	*<0.001*
Daily activity (≥8.25 METs·h/day)	1.033	0.493 to 1.574	*<0.001*
Habitual exercise (≥0.73 METs·h/day)	1.480	0.938 to 2.021	*<0.001*
Egg intake (≥3 times/week)	0.856	0.270 to 1.442	*0.004*
BMI (≥23.0)	-8.738	-9.283 to -8.194	*<0.001*
Age (≥57 years)	-1.496	-2.040 to -0.952	*<0.001*
Sex (women)	9.567	8.931 to 10.203	*<0.001*

BMI, body mass index; CI, confidence interval; Coeff., coefficient; HDL-C, high density lipoprotein cholesterol; METs, metabolic equivalents.

[a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake, and BMI. The coefficient value represents change in HDL-C per dummy variable (0, 1) of environmental factors.

4.2 Analysis between HDL-C levels and related SNPs

The 7 major SNPs selected from the 65 SNPs in the GWAS catalog according to the gene and cytoBand groups were rs3764261 in *HERPUD1–CETP* (16q13), rs662799 in *APOA5* (11q23.3), rs1800588 in *LIPC* (15q21.3), rs328 in *LPL* (8p21.3), rs2575876 in *ABCA1* (9q31.1), rs3786247 in *LIPG* (18q21.1), and rs429358 in *APOE* (19q13.32) (Table 4). The frequencies (0.100 to 0.649) and coefficients (-4.003 to 8.863) varied for each SNP, and the highest coefficient was observed for rs3764261.

Table 4. Multivariate regression analysis between HDL-C levels and seven HDL-C related SNPs from the GWAS catalog

SNP	cytoBand	REF/ALT	Gene	Frequency of ALT	Coeff.[a]	95% CI	*P*-value
rs3764261	16q13	C/A	*HERPUD1, CETP*	0.207	8.863	7.958 to 9.770	6.07×10^{-82}
rs662799	11q23.3	G/A	*APOA5*	0.649	5.713	4.932 to 6.494	1.12×10^{-46}
rs1800588	15q21.3	C/T	*LIPC*	0.510	4.447	3.700 to 5.194	1.76×10^{-31}
rs328	8p21.3	C/G	*LPL*	0.126	6.136	5.006 to 7.266	1.77×10^{-26}
rs2575876	9q31.1	G/A	*ABCA1*	0.276	-4.003	-4.840 to -3.164	7.67×10^{-21}
rs3786247	18q21.1	T/G	*LIPG*	0.460	3.209	2.452 to 3.966	1.02×10^{-16}
rs429358	19q13.32	T/C	*APOE*	0.100	-3.594	-4.864 to -2.322	2.89×10^{-8}

ALT, alternative allele; BMI, body mass index; CI, confidence interval; Coeff., coefficient; GWAS, genome-wide association study; HDL-C, high density lipoprotein cholesterol; REF, referent allele; SNP, single nucleotide polymorphism.

[a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake and BMI. The coefficient value represents change in HDL-C per ALT allele copy (0, 1, 2) for the SNP.

4.3 Interaction between HDL-C levels to smoking, drinking and related SNPs

The HDL-C levels varied for each genotype group based on the smoking and drinking status (Table 5). The highest HDL-C level (74.6 mg/dL, 70.8 - 78.4) was observed in heavy drinkers with the rs3764261 alternative homo-genotype (AA), while the lowest was observed in current smokers with the rs662799 referent homo-genotype (GG) and hetero-genotype (GA). The gene-environment interactions between the 7 SNPs and smoking/drinking were not statistically significant, and the lowest p-value of 0.004 was higher than the p-value obtained after applying Bonferroni correction ($P < 0.00077$). These interactions were not statistically significant for all 65 SNPs selected from the GWAS catalog (Table A.1). No significant interaction was observed in the subgroup analysis based on sex (data not shown in Table A.1).

Table 5. Interaction between HDL-C levels according to different smoking and drinking statues and the 7 selected SNPs

	Smoking					Drinking				
	Never & former		Current		*P*-value for interaction[a]	Non-moderate		Heavy[b]		*P*-value for interaction[a]
	RR & RA	AA	RR & RA	AA		RR & RA	AA	RR & RA	AA	
	HDL-C	HDL-C	HDL-C	HDL-C		HDL-C	HDL-C	HDL-C	HDL-C	
rs3764261	64.0	72.2	56.0	64.7	0.156	62.5	69.5	62.2	74.6	0.015
	(63.7 - 64.3)	(70.6 - 73.8)	(55.4 - 56.6)	(60.7 - 68.8)		(62.1 - 62.8)	(67.9 - 71.2)	(61.6 - 62.9)	(70.8 - 78.4)	
	N=8,723	N=408	N=2,130	N=109		N=8,464	N=405	N=2,389	N=112	
rs662799	62.9	66.5	54.7	58.5	0.706	61.5	64.6	60.7	65.6	0.004
	(62.4 - 63.3)	(66.0 - 67.0)	(53.9 - 55.5)	(57.5 - 59.5)		(61.0 - 61.9)	(64.1 - 65.1)	(59.9 - 61.6)	(64.5 - 66.6)	
	N=5,301	N=3830	N=1,236	N=1003		N=5,094	N=3775	N=1,443	N=1058	
rs1800588	63.6	66.6	55.3	59.5	0.312	62.0	65.0	61.7	65.8	0.387
	(63.2 - 64.0)	(65.9 - 67.2)	(54.6 - 56.0)	(58.2 - 60.8)		(61.6 - 62.4)	(64.4 - 65.7)	(61.0 - 62.5)	(64.5 - 67.2)	
	N=6,723	N=2408	N=1,661	N=578		N=6,530	N=2339	N=1,854	N=647	
rs328	64.3	70.4	56.4	58.4	0.735	62.7	69.0	62.8	64.8	0.658
	(63.9 - 64.6)	(67.6 - 73.2)	(55.7 - 57.0)	(52.3 - 64.4)		(62.4 - 63.0)	(66.1 - 72.0)	(62.1 - 63.4)	(59.6 - 70.1)	
	N=8,983	N=148	N=2,207	N=32		N=8,721	N=148	N=2,469	N=32	
rs2575876	64.6	62.1	56.6	54.6	0.476	63.0	60.5	63.0	61.0	0.354
	(64.2 - 64.9)	(61.0 - 63.3)	(55.9 - 57.2)	(52.4 - 56.7)		(62.6 - 63.3)	(59.3 - 61.6)	(62.3 - 63.6)	(58.6 - 63.4)	
	N=8,436	N=695	N=2,062	N=177		N=8,207	N=662	N=2,291	N=210	
rs3786247	63.9	66.0	55.6	59.3	0.670	62.3	64.6	62.2	65.0	0.569
	(63.6 - 64.3)	(65.3 - 66.7)	(55.0 - 56.3)	(57.8 - 60.7)		(62.0 - 62.7)	(63.8 - 65.3)	(61.5 - 62.9)	(63.5 - 66.5)	
	N=7,232	N=1899	N=1,770	N=469		N=7,029	N=1840	N=1,973	N=528	
rs429358	64.4	61.6	56.4	57.9	0.931	62.8	60.7	62.8	61.1	0.723
	(64.1 - 64.7)	(58.1 - 65.0)	(55.8 - 57.0)	(50.1 - 65.6)		(62.5 - 63.2)	(57.4 - 64.0)	(62.1 - 63.5)	(52.3 - 70.0)	
	N=9,037	N=94	N=2,214	N=25		N=8,773	N=96	N=2,478	N=23	

AA, alterative homo-genotype; BMI, body mass index; HDL-C, high density lipoprotein cholesterol; RA, referent and alterative allele hetero-genotype; RR, referent allele homo-genotype; SNP, single nucleotide polymorphism. [a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake, and BMI. [b] ≥20 g alcohol/day.

4.4 Population attributable fractions of non- and genetic factors for low HDL-C

The ORs for low HDL-C were statistically significant for several non-genetic factors, including smoking, drinking, BMI, age and sex, and for the genetic factors, 6 SNPs except rs1800588 (Table 6). The PAF for low HDL-C in the non-genetic factors was the highest in men (63.2%), and the PAFs of smoking and drinking were 23.1% and 41.8%, respectively. The PAF for low HDL-C in the genetic factors was the highest in rs3764261 (31.5%), which was higher than that of smoking and lower than that of drinking. The impact of the PAFs of three SNPs (25.5%) and 4-6 SNPs (23.7%) according to the number of SNPs with high-risk genotype for low HDL-C was similar to that of smoking, although the ORs for low HDL-C showed an apparent increasing trend with the number of SNPs with higher-risk genotype ($P < 0.001$).

Table 6. Population attributable fractions of non-genetic and genetic factors for low HDL-C

	Proportion of exposure	OR[a]	95% CI	PAF (%)
Non-genetic factors				
Smoking habit (current)	41.8	2.23	1.85-2.70	23.1
Drinking habit (< 20 gram alcohol/day)	76.8	2.19	1.77-2.71	41.8
Daily activity (< 8.25 METs/day)	62.7	1.11	0.93-1.33	-
Habitual exercise (< 0.73 METs/day)	52.4	1.14	0.95-1.36	-
Egg intake (< 3 times/week)	70.4	0.93	0.77-1.12	-
BMI (≥ 23.0)	72.6	2.35	1.93-2.85	41.6
Age (≥57 years)	54.5	1.44	1.20-1.72	16.6
Sex (men)	80.4	4.68	3.72-5.89	63.2
Genetic factors				
rs3764261 (RR)	73.5	1.75	1.44-2.13	31.5
rs662799 (RR)	26.6	2.89	2.35-3.55	17.4
rs1800588 (RR)	27.3	1.16	0.95-1.41	-
rs328 (RR)	81.1	1.36	1.09-1.70	21.6
rs2575876 (RA & AA)	55.6	1.43	1.20-1.71	16.8
rs3786247 (RR)	34.5	1.36	1.13-1.64	9.2
rs429358 (RA & AA)	24.6	1.56	1.27-1.92	8.9
Number of SNPs with high-risk genotype[b]				
0-1 SNPs	7.3	1.00	-	-
2 SNPs	26.4	1.97	1.38-2.82	13.0
3 SNPs	37.4	3.16	2.24-4.47	25.5
4-6 SNPs	29.0	5.49	3.84-7.84	23.7
P for trend		*<0.001*		

AA, alterative allele homo-genotype; BMI, body mass index; CI, confidence interval; HDL-C, high density

a) Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake and BMI.

b) SNP of rs1800588 is excluded.

4.5 Population-based impact of smoking, drinking, and genetic factors on whole levels of HDL-C

To estimate the another population-based impact of the combination of SNPs, the cutoff value for the population-based prevalence at an exposure level similar to that for environmental factors was fixed at 21.3% for the combinations of 7 and 65 SNPs, too. At 21.3% prevalence, the coefficient of the combination of 7 SNPs with the HDL-C levels was 8.012 (95% CI, 7.379 to 8.646, $P < 0.001$), which was 1.07 times greater than that for drinking (7.498, 6.791 to 8.205, 22.0% at prevalence), and 1.46 times greater than that for smoking (-5.477, -6.184 to -4.770, 19.7%) (Table 7). 1The interaction between the combination of the 7 SNPs and drinking was statistically significant with a positive direction (2.647, 1.087 to 4.207, $P = 0.001$), whereas that between the combination of the 7 SNPs and smoking was not ($P = 0.993$). At 21.2% prevalence, the coefficient of the combination of the 65 SNPs with the HDL-C levels was 7.109 (95% CI, 6.460 to 7.759, $P < 0.001$), which was 0.89 times lesser than that for the combination of the 7 SNPs; however, the difference was not statistically significant. The interaction between the combination of the 65 SNPs and drinking was also statistically significant (2.643, 1.085 to 4.200, $P < 0.001$).

Table 7. Population-based impact of smoking, drinking, and combined genetic factors on whole levels of HDL-C at similar prevalence

	N	Prevalence (%)	Coeff.[a]	95% CI	*P*-value	Interaction		
						Coeff.[a]	95% CI	*P*-value
Smoking (current)	2,239	19.7	-5.477	-6.184 to -4.770	<0.001	0.007	-1.576 to 1.587	0.993
Drinking (≥20 g alcohol/day)	2,501	22.0	7.498	6.791 to 8.205	<0.001	2.647	1.087 to 4.207	0.001
Combination of 7 SNPs	2,418	21.3	8.012	7.379 to 8.646	<0.001	-	-	-
Smoking (current)	2,239	19.7	-5.422	-6.133 to -4.710	<0.001	0.900	-0.706 to 2.505	0.319
Drinking (≥20 g alcohol/day)	2,501	22.0	7.324	6.613 to 8.036	<0.001	2.643	1.085 to 4.200	<0.001
Combination of 65 SNPs	2,418	21.3	7.109	6.460 to 7.759	<0.001	-	-	-

HDL-C, high density lipoprotein cholesterol; Coeff., coefficient; CI, confidence interval; SNPs, single nucleotide polymorphisms; BMI, body mass index.

[a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake BMI and combination of 7/65 SNPs. The coefficient value represents change in HDL-C per dummy variable (0, 1) of smoking, drinking, and combined genetic factor.

Sensitivity analysis of the coefficients of drinking and the combinations of 7 and 65 SNPs showed the dose-response relationship based on three levels of exposure at 10.1% and 10.0%, 21.3%, and 40.9% prevalence (Table 8).

Table 8. Sensitivity analysis of the coefficients for association of HDL-C levels with drinking and the combined SNPs at different prevalence

	Prevalence	Coeff.[a]	95% CI	*P*-value
Low prevalence				
Drinking (≥ 46 g	10.2	7.617	6.688 to 8.563	*<0.001*
Combination of 7 SNPs	10.1	8.862	7.988 to 9.737	*<0.001*
Combination of 65 SNPs	10.0	8.207	7.332 to 9.082	*<0.001*
Moderate prevalence				
Drinking (≥ 20 g	22.0	7.498	6.791 to 8.205	*<0.001*
Combination of 7 SNPs	21.3	8.012	7.379 to 8.646	*<0.001*
Combination of 65 SNPs	21.3	7.109	6.460 to 7.759	*<0.001*
High prevalence				
Drinking (≥ 4.7 g	40.1	5.914	5.307 to 6.522	*<0.001*
Combination of 7 SNPs	40.9	6.795	6.268 to 7.323	*<0.001*
Combination of 65 SNPs	40.9	6.308	5.778 to 6.837	*<0.001*

HDL-C, high density lipoprotein cholesterol; Coeff., coefficient; CI, confidential interval; METs, metabolic equivalents; BMI, body mass index.

[a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake, and BMI. The combination of 7 SNPs was included as an adjusting variable for drinking.

CHAPTER 5: DISCUSSION

In the present study, we observed significant associations between HDL-C levels and smoking, drinking, daily activity, habitual exercise, egg intake, BMI, age, sex and 7 SNPs in *CETP*, *APOA5*, *LIPC*, *LPL*, *ABCA1*, *LIPG*, and *APOE*. The PAFs, as a population-based impact, for low HDL-C were the highest in men on the non-genetic factors and in *CETP* rs3764261 on the genetic factors. The impact of the genetic factor PAF was higher than that of smoking and was lower than that of drinking.

Genetic factors that affect HDL-C levels, such as SNPs, are primarily associated with genes that encode enzymes from the RCT system, such as *ABCA1*, *LCAT, CETP*, *LIPC*, *APOA1/C3/A4/A5*, *SCARB1*, and *LPL*.[2, 7] The SNPs in the corresponding genes, except those in *LCAT* and *SCARB1*, were considered among the 7 major SNPs selected in the present analysis. The SNPs in *SCARB1* were not included because the two SNPs with genome-wide significance were not listed in the GWAS catalog, and the lowest p-value for the *SCARB1* SNP (rs838886) listed in the catalog was higher than the genome-wide significance ($P = 7.34 \times 10^{-8}$; data not shown in Table A.1). As the MAF of *LCAT* was less than 0.01, the SNPs of

LCAT were excluded from the GWAS analysis. The SNPs in *LIPG* and *APOE*, which are associated with HDL-C production via a system different from RCT, were also considered among the 7 major SNPs.[10,12] The genetic variants of *CETP* were observed to exhibit the most significant influence on HDL-C levels, which was concordant with findings from previous reports.[8-10]

Cigarette smoking is associated with lower HDL-C levels, even though the mechanisms are yet to be completely elucidated. Certain studies have shown that smoking is related to ApoA1 concentration[13] and CETP activity[14]; however, these results could be considered controversial.[70,71] Alcohol consumption is reported to be associated with increased expression of ABCA1[72] and a higher APOA1 concentration[73] in peripheral blood and a lower CETP activity.[74]

In the present study, the interaction of the 65 and 7 SNPs with drinking was not statistically significant after Bonferroni correction was applied. Previous studies reported significant association of alcohol consumption and polymorphisms in multiple genes (*CETP*, *APOA1*/*A2*, *LPL*, *ADH3*, *ADH1*, and *ALDH2*) with HDL-C levels.[45,75-78] The association between CETP and ABCA1 expressions, and alcohol consumption has been also reported in previous studies, but their mechanism is not clear.[72,74] However, no genome-wide significance was reported in the gene-alcohol

interaction for *CETP*, *APOA5*, *LIPC*, and *LPL* in a particular GWAS.[79] The interaction between each SNP and smoking was not statistically significant after Bonferroni correction was applied, too. These results suggest that genetic factors may have a minor or minimal impact on the interaction with drinking and smoking.

Several studies have previously reported the association between SNPs and HDL-C levels, which have been listed in the GWAS catalog. In the present study, we selected the 498 SNPs listed in the GWAS results that were a part of the J-MICC Study, and observed 65 SNPs with genome-wide significance for the analysis. We selected 7 SNPs according to the gene and cytoBand groups. The Manhattan plot for total SNPs consistently showed seven peaks, except that for *SCARB1*. These observations support proposition that the 7 SNPs are appropriate representatives of the SNPs associated with HDL-C levels in the present analysis.

In the present study, we investigated the population-based impact of both non-genetic and genetic factors on low HDL-C, using PAF. The OR for low HDL-C was used as the relative risk when the PAF was calculated, because the prevalence of low HDL-C was obtained from the baseline general population and its rate was relatively low (5.0% in both sexes).[80,81] To the best of our knowledge, studies investing the PAF for low HDL-C with non-genetic and/or genomic factors have

not yet been conducted. The highest PAFs was observed in men on the non-genetic factors and in *CETP* rs3764261 on the genetic factors. The impact of the genetic factor PAF was higher than that of smoking and was lower than that of drinking. These observations suggest that, from a public health perspective, the population-based impact of genomic factors for low HDL-C is comparably high compared to non-genetic factors.

The present study also demonstrated that another population-based impact of the combination of the 7 representative SNPs on whole levels of HDL-C without clinical cutoff point (<40 mg/dL) was similar to that of drinking, and was 1.46 time higher than that of smoking. The impact of the combination of the 65 SNPs was 0.87 times lower than that of the combination of the 7 SNPs. The PAF on low HDL-C may have more important impact for disease burden and its prevention, and this another population-based impact on whole levels of HDL-C without clinical cutoff point may reflect a biological effect on whole levels.

The strength of this study is that the population-based impact of non-genetic and genetic factors on HDL-C levels was evaluated simultaneously using data from an adequate number of subjects and total gene information. To our knowledge, this

is the first comprehensive report on the population-based impact of the abovementioned factors.

Meanwhile, the present study has several limitations. First, a causal relationship was not confirmed, as this is a cross-sectional study. Second, atheroprotective and non-atheroprotective HDL particles were jointly considered as total HDL-C. The two fractions of HDL2-C and HDL3-C have different effects on CVD risk.2 Third, the present study selected 7 representative SNPs to estimate the population-based impact; the highest impact may have been estimated because the highest coefficients of the 7 representative SNPs were selected based on the gene and cytoBand groups. Fourth, the replication test on GWAS was not conducted, because the present study used information from the GWAS catalog in which the association between SNPs and HDL-C levels had been estimated and published previously. Fifth, the effect of residual SNPs (those apart from the 65 SNPs), referred to as "missing heritability", was not considered. The polygenic risk score may support the estimation of this effect.[66] Sixth, PAF valid only in the absence of confounding and/or effect modification.[80] The lack of unknown data on confounding is likely to misestimate the true PAF, the extent to which is dependent on the magnitude of confounding.[81] Furthermore, PAF estimate is restricted by time

and population and depends on the quality and representativeness of the exposure and risk data.

In conclusion, the present study demonstrated that the population-based impact of genomic factor *CETP* rs3764261 for low HDL-C was higher than that of smoking and lower than that of drinking.

REFERENCES

1. National Cholesterol Education Program . Expert Panel on D, Treatment of High Blood Cholesterol in A. Third report of the National Cholesterol Education Program (NCEP) Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III): The Program; 2002. DOI.

2. von Eckardstein A, Kardassis D. High density lipoproteins: from biological understanding to clinical exploitation: Springer Nature; 2015. DOI.

3. Volcik K, Ballantyne CM, Pownall HJ, Sharrett AR, Boerwinkle E. Interaction effects of high-density lipoprotein metabolism gene variation and alcohol consumption on coronary heart disease risk: the atherosclerosis risk in communities study. Journal of studies on alcohol and drugs. 2007;68(4):485-92.

4. Wada M, Daimon M, Emi M, Iijima H, Sato H, Koyano S, et al. Genetic association between aldehyde dehydrogenase 2 (ALDH2) variation and high-density lipoprotein cholesterol (HDL-C) among non-drinkers in two large population samples in Japan. Journal of atherosclerosis and thrombosis. 2008;15(4):179-84.

5. Campbell SC, Moffatt RJ, Stamford BA. Smoking and smoking cessation—the relationship between cardiovascular disease and lipoprotein

metabolism: a review. Atherosclerosis. 2008;201(2):225-35.

6. He Bm, Zhao Sp, Peng Zy. Effects of cigarette smoking on HDL quantity and function: implications for atherosclerosis. Journal of cellular biochemistry. 2013;114(11):2431-6.

7. Weissglas-Volkov D, Pajukanta P. Genetic causes of high and low serum HDL-cholesterol. Journal of lipid research. 2010;51(8):2032-57.

8. Hiura Y, Shen C-S, Kokubo Y, Okamura T, Morisaki T, Tomoike H, et al. Identification of Genetic Markers Associated With High-Density Lipoprotein-Cholesterol by Genome-Wide Screening in a Japanese Population The Suita Study. Circulation Journal. 2009;73(6):1119-26.

9. Nakamura A, Niimura H, Kuwabara K, Takezaki T, Morita E, Wakai K, et al. Gene-gene combination effect and interactions among ABCA1, APOA1, SR-B1, and CETP polymorphisms for serum high-density lipoprotein-cholesterol in the Japanese population. PloS one. 2013;8(12):e82046.

10. Heid IM, Boes E, Müller M, Kollerits B, Lamina C, Coassin S, et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. Circulation: Cardiovascular Genetics. 2008;1(1):10-20.

11. Thompson JF, Wood LS, Pickering EH, DeChairo B, Hyde CL. High-density genotyping and functional SNP localization in the CETP gene. Journal of lipid research. 2007;48(2):434-43.

12. Surakka I, Horikoshi M, Mägi R, Sarin A-P, Mahajan A, Lagou V, et al. The impact of low-frequency and rare variants on lipid levels. Nature genetics. 2015;47(6):589-97.

13. Sigurdsson Jr G, Gudnason V, Sigurdsson G, Humphries SE. Interaction between a polymorphism of the apo AI promoter region and smoking determines plasma levels of HDL and apo AI. Arteriosclerosis and thrombosis: a journal of vascular biology. 1992;12(9):1017-22.

14. Dullaart RP, Hoogenberg K, Dikkeschei BD, van Tol A. Higher plasma lipid transfer protein activities and unfavorable lipoprotein changes in cigarette-smoking men. Arteriosclerosis and thrombosis: a journal of vascular biology. 1994;14(10):1581-5.

15. Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. British medical journal. 1989;298(6676):784-8.

16. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS,

et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences. 2009;106(23):9362-7.

17. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic acids research. 2014;42(D1):D1001-D6.

18. Kontush A, Lindahl M, Lhomme M, Calabresi L, Chapman MJ, Davidson WS. Structure of HDL: particle subclasses and molecular components. High Density Lipoproteins. 2015:3-51.

19. Hewing B, Moore KJ, Fisher EA. HDL and cardiovascular risk: time to call the plumber? Circ Res. 2012;111(9):1117-20.

20. Güleç SEC, Erol C. High-density lipoprotein cholesterol and risk of cardiovascular disease. J Cardiol Pract. 2020;19:133-4.

21. Rothblat GH, Phillips MC. High-density lipoprotein heterogeneity and function in reverse cholesterol transport. Curr Opin Lipidol. 2010;21(3):229-38

22. Kontush A, Lhomme M, Chapman MJ. Unraveling the complexities of the HDL lipidome1. Journal of lipid research. 2013;54(11):2950-63.

23. Rader DJ, Hobbs HH. Disorders of lipoprotein metabolism. Harrisons

principles of internal medicine. 2005;16(2):2286.

24. Timmins JM, Lee J-Y, Boudyguina E, Kluckman KD, Brunham LR, Mulya A, et al. Targeted inactivation of hepatic Abca1 causes profound hypoalphalipoproteinemia and kidney hypercatabolism of apoA-I. The Journal of clinical investigation. 2005;115(5):1333-42.

25. Brunham LR, Kruit JK, Iqbal J, Fievet C, Timmins JM, Pape TD, et al. Intestinal ABCA1 directly contributes to HDL biogenesis in vivo. The Journal of clinical investigation. 2006;116(4):1052-62.

26. Favari E, Chroni A, Tietge UJF, Zanotti I, Bernini F. Cholesterol efflux and reverse cholesterol transport. High density lipoproteins. 2015:181-206

27. Zannis VI, Chroni A, Krieger M. Role of apoA-I, ABCA1, LCAT, and SR-BI in the biogenesis of HDL. Journal of molecular medicine. 2006;84(4):276-94.

28. Nazir S, Jankowski V, Bender G, Zewinger S, Rye K-A, van der Vorst EPC. Interaction between high-density lipoproteins and inflammation: Function matters more than concentration! Advanced Drug Delivery Reviews. 2020;159:94-119.

29. Sundjaja JH, Pandey S. Cholesterol Screening. StatPearls [Internet]. Last Update: May 10, 2021.

30. Félix-Redondo FJ, Grau M, Fernández-Bergés D. Cholesterol and

cardiovascular disease in the elderly. Facts and gaps. Aging Dis. 2013;4(3):154-69.

31. Okada M, Matsui H, Ito Y, Fujiwara A. Direct measurement of HDL cholesterol: method eliminating apolipoprotein E-rich particles. Journal of clinical laboratory analysis. 2001;15(4):223-9.

32. Hafiane A, Genest J. High density lipoproteins: measurement techniques and potential biomarkers of cardiovascular risk. BBA clinical. 2015;3:175-88.

33. on Smoking O, Centers for Disease C, Prevention. Chemistry and Toxicology of Cigarette Smoke and Biomarkers of Exposure and Harm. How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General: Centers for Disease Control and Prevention (US); 2010.

34. West R. Tobacco smoking: Health impact, prevalence, correlates and interventions. Psychol Health. 2017;32(8):1018-36

35. West R, McEwen A, Bolling K, Owen L. Smoking cessation and smoking patterns in the general population: a 1-year follow-up. Addiction. 2001;96(6):891-902

36. Cao S, Yang C, Gan Y, Lu Z. The health effects of passive smoking: an overview of systematic reviews based on observational epidemiological evidence.

PloS one. 2015;10(10):e0139907

37. Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), Office on Smoking and Health (US). How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the surgeon general. 2010

38. Iarc I. Monograph on the Evaluation of Carcinogenic Risks to Humans, Vol 44. Alcohol Drinking. IARC, Lyon; 1988.

39. Rehm J. The risks associated with alcohol use and alcoholism. Alcohol Res Health. 2011;34(2):135-43

40. Gunzerath L, Faden V, Zakhari S, Warren K. National Institute on Alcohol Abuse and Alcoholism report on moderate drinking. Alcoholism: Clinical and experimental research. 2004;28(6):829-47

41. Katzung BG, Masters SB, Trevor AJ. Basic and Clinical Pharmacology. McGraw-Hill Companies, Inc. 2012

42. Pavlova SI, Jin L, Gasparovich SR, Tao L. Multiple alcohol dehydrogenases but no functional acetaldehyde dehydrogenase causing excessive acetaldehyde production from ethanol by oral streptococci. Microbiology. 2013;159(Pt 7):1437

43. Beier JI, McClain CJ. Mechanisms and cell signaling in alcoholic liver disease. 2010

44. Brunton LL, Chabner B, Knollmann BC. Goodman & Gilman's the pharmacological basis of therapeutics: McGraw-Hill Education New York, NY, USA:; 2018. DOI.

45. De Oliveira e Silva ER, Foster D, McGee Harper M, Seidman CE, Smith JD, Breslow JL, et al. Alcohol consumption raises HDL cholesterol levels by increasing the transport rate of apolipoproteins AI and A-II. Circulation. 2000;102(19):2347-52

46. Van der Gaag MS, Van Tol A, Vermunt SHF, Scheek LM, Schaafsma G, Hendriks HFJ. Alcohol consumption stimulates early steps in reverse cholesterol transport. Journal of lipid research. 2001;42(12):2077-83

47. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. Nature reviews Genetics. 2005;6(2):109-18

48. Scherer SW, Visscher PM. Genome-wide association studies: From polymorphism to personalized medicine: Cambridge University Press; 2016. DOI.

49. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS

Comput Biol. 2012;8(12):e1002822-e

50. Stram DO. Design, analysis, and interpretation of genome-wide association scans: Springer; 2014. DOI.

51. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS computational biology. 2012;8(12):e1002822

52. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nature reviews Genetics. 2010;11(5):356-66

53. Ordovas JM, Robertson R, Cleirigh EN. Gene-gene and gene-environment interactions defining lipid-related traits. Current opinion in lipidology. 2011;22(2):129-36

54. Hiura Y, Shen CS, Kokubo Y, Okamura T, Morisaki T, Tomoike H, et al. Identification of genetic markers associated with high-density lipoprotein-cholesterol by genome-wide screening in a Japanese population: the Suita study. Circulation journal : official journal of the Japanese Circulation Society. 2009;73(6):1119-26

55. Weissglas-Volkov D, Pajukanta P. Genetic causes of high and low serum HDL-cholesterol. Journal of lipid research. 2010;51(8):2032-57

56. Nakamura A, Niimura H, Kuwabara K, Takezaki T, Morita E, Wakai K, et al. Gene-gene combination effect and interactions among ABCA1, APOA1, SR-B1, and CETP polymorphisms for serum high-density lipoprotein-cholesterol in the Japanese population. PloS one. 2013;8(12):e82046

57. Cavalli M, Pan G, Nord H, Wadelius C. Looking beyond GWAS: allele-specific transcription factor binding drives the association of GALNT2 to HDL-C plasma levels. Lipids in health and disease. 2016;15:18

58. Heid IM, Boes E, Muller M, Kollerits B, Lamina C, Coassin S, et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. Circulation Cardiovascular genetics. 2008;1(1):10-20

59. Hamajima N. The Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study) to detect gene-environment interactions for cancer. Asian Pacific journal of cancer prevention : APJCP. 2007;8(2):317-23

60. Wakai K, Hamajima N, Okada R, Naito M, Morita E, Hishida A, et al. Profile of participants and genotype distributions of 108 polymorphisms in a cross-sectional study of associations of genotypes with lifestyle and clinical factors: a project in the Japan Multi-Institutional Collaborative Cohort (J-MICC) Study.

Journal of epidemiology. 2011;21(3):223-35

61. Hamajima N, Wakai K, Naito M, Nishio K, Ishida Y, Okada R, et al. The Japan Multi-institutional Collaborative Cohort Study (JMICC Study) to detect gene-environment interactions for cancer. Asian Pacific Journal of Cancer Prevention. 2007;8(2):317-23

62. Naito M, Eguchi H, Okada R, Ishida Y, Nishio K, Hishida A, et al. Controls for monitoring the deterioration of stored blood samples in the Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study). Nagoya journal of medical science. 2008;70(3-4):107-15

63. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):s13742-015

64. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. The American Journal of Human Genetics. 2008;83(4):445-56

65. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS genetics. 2006;2(12):e190

66. Genomes Project C. A global reference for human genetic variation. Nature. 2015;526(7571):68

67. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nature methods. 2013;10(1):5-6

68. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nature genetics. 2016;48(10):1284-7

69. Glass S, Dwyer GB, American College of Sports M. ACSM'S metabolic calculations handbook: Lippincott Williams & Wilkins; 2007. DOI.

70. Lesna IK, Poledne R, Pagacova L, Stavek P, Pitha J. HDL and apolipoprotein A1 concentrations as markers of cholesterol efflux in middle-aged women: interaction with smoking. Neuro Endocrinol Lett. 2012;33(Suppl 2):38-42

71. Freeman DJ, Caslake MJ, Griffin BA, Hinnie J, Tan CE, Watson TD, et al. The effect of smoking on post-heparin lipoprotein and hepatic lipase, cholesteryl ester transfer protein and lecithin: cholesterol acyl transferase activities in human plasma. European journal of clinical investigation. 1998;28(7):584-91

72. Hoang A, Tefft C, Duffy SJ, Formosa M, Henstridge DC, Kingwell BA, et

al. ABCA1 expression in humans is associated with physical activity and alcohol consumption. Atherosclerosis. 2008;197(1):197-203

73. Rimm EB, Williams P, Fosher K, Criqui M, Stampfer MJ. Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. Bmj. 1999;319(7224):1523-8

74. Hannuksela M, Marcel YL, Kesäniemi YA, Savolainen MJ. Reduction in the concentration and activity of plasma cholesteryl ester transfer protein by alcohol. Journal of lipid research. 1992;33(5):737-44

75. Mehlig K, Strandhagen E, Svensson P-A, Rosengren A, Torén K, Thelle DS, et al. CETP TaqIB genotype modifies the association between alcohol and coronary heart disease: the INTERGENE case-control study. Alcohol. 2014;48(7):695-700

76. Baik I, Lee S, Kim SH, Shin C. A lipoprotein lipase gene polymorphism interacts with consumption of alcohol and unsaturated fat to modulate serum HDL-cholesterol concentrations. The Journal of nutrition. 2013;143(10):1618-25

77. Hines LM, Stampfer MJ, Ma J, Gaziano JM, Ridker PM, Hankinson SE, et al. Genetic variation in alcohol dehydrogenase and the beneficial effect of moderate alcohol consumption on myocardial infarction. New England Journal of

Medicine. 2001;344(8):549-55

78. Yokoyama A, Yokoyama T, Matsui T, Mizukami T, Kimura M, Matsushita S, et al. Alcohol dehydrogenase-1B (rs1229984) and aldehyde dehydrogenase-2 (rs671) genotypes are strong determinants of the serum triglyceride and cholesterol levels of Japanese alcoholic men. PLoS One. 2015;10(8):e0133460

79. Marques-Vidal P, Bochud M, Paccaud F, Waterworth D, Bergmann S, Preisig M, et al. No interaction between alcohol consumption and HDL-related genes on HDL cholesterol levels. Atherosclerosis. 2010;211(2):551-7

80. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. American journal of public health. 1998;88(1):15-9

81. Shield KD, Parkin DM, Whiteman DC, Rehm J, Viallon V, Micallef CM, et al. Population attributable and preventable fractions: cancer risk factor surveillance, and cancer policy projection. Current epidemiology reports. 2016;3(3):201-11

Table A.1. Sixty-five HDL-C related SNPs selected from the GWAS catalog, and their association with HDL-C levels and interaction with smoking and drinking

SNPs	cytoBand	REF/ ALT	Gene	Frequency of ALT	Association with		Interaction with					
					HDL-C levels		smoking			drinking		
					Coeff.[a]	*P*-value	Coeff.[a]	SE	*P*-value	Coeff.[a]	SE	*P*-value
rs328	8p21.3	C/G	*LPL*	0.126	6.136	1.77×10^{-26}	-0.247	0.730	0.735	-0.310	0.701	0.658
rs325	8p21.3	T/C	*LPL*	0.126	6.131	2.03×10^{-26}	-0.249	0.731	0.734	-0.296	0.701	0.673
rs79407615	8p21.3	T/G	*LPL, SLC18A1*	0.125	6.141	2.18×10^{-26}	-0.135	0.733	0.854	-0.133	0.704	0.85
rs12678919	8p21.3	A/G	*LPL, SLC18A1*	0.125	6.131	2.21×10^{-26}	-0.081	0.732	0.912	-0.168	0.704	0.811
rs10503669	8p21.3	C/A	*LPL, SLC18A1*	0.125	6.129	2.69×10^{-26}	-0.128	0.733	0.861	-0.127	0.704	0.857
rs7841189	8p21.3	C/T	*LPL, SLC18A1*	0.125	6.105	4.27×10^{-26}	-0.113	0.733	0.877	-0.186	0.704	0.792
rs10096633	8p21.3	C/T	*LPL, SLC18A1*	0.125	6.084	6.24×10^{-26}	-0.126	0.733	0.863	-0.201	0.704	0.775
rs17482753	8p21.3	G/T	*LPL, SLC18A1*	0.125	6.071	7.96×10^{-26}	-0.118	0.733	0.872	-0.193	0.704	0.784
rs13702	8p21.3	T/C	*LPL*	0.186	5.087	3.23×10^{-25}	0.628	0.624	0.314	0.209	0.596	0.726
rs15285	8p21.3	C/T	*LPL*	0.186	5.087	3.23×10^{-25}	0.628	0.624	0.314	0.209	0.596	0.726

rs287	8p21.3	A/G	*LPL*	0.192	5.141	4.56×10^{-25}	0.434	0.632	0.493	0.062	0.604	0.918
rs2083637	8p21.3	A/G	*LPL, SLC18A1*	0.186	5.021	1.19×10^{-24}	0.630	0.623	0.312	0.199	0.597	0.739
rs326	8p21.3	A/G	*LPL*	0.188	4.981	2.15×10^{-24}	0.576	0.621	0.354	0.196	0.594	0.741
rs11984636	8p21.3	T/C	*LPL, SLC18A1*	0.118	5.761	2.01×10^{-22}	0.195	0.748	0.794	-0.174	0.714	0.808
rs115849089	8p21.3	G/A	*LPL, SLC18A1*	0.130	5.313	1.35×10^{-19}	0.409	0.750	0.586	0.111	0.714	0.877
rs9644568	8p21.3	G/A	*LPL, SLC18A1*	0.128	5.673	1.16×10^{-18}	0.356	0.827	0.667	0.092	0.780	0.906
rs4244457	8p21.3	C/T	*LPL, SLC18A1*	0.268	3.305	1.84×10^{-14}	0.761	0.557	0.172	0.619	0.526	0.239
rs28526159	8p21.3	C/T	*LPL, SLC18A1*	0.274	3.075	7.75×10^{-13}	0.760	0.553	0.169	0.439	0.523	0.401
rs2575876	9q31.1	G/A	*ABCA1*	0.276	-4.003	7.67×10^{-21}	0.381	0.535	0.476	0.472	0.509	0.354
rs1883025	9q31.1	C/T	*ABCA1*	0.283	-3.97	8.59×10^{-21}	0.384	0.531	0.470	0.648	0.506	0.201
rs2740488	9q31.1	A/C	*ABCA1*	0.279	-3.904	5.62×10^{-20}	0.386	0.534	0.470	0.376	0.509	0.46
rs12686004	9q31.1	G/A	*ABCA1*	0.234	-3.409	3.16×10^{-14}	0.247	0.564	0.661	0.561	0.534	0.293
rs4149268	9q31.1	C/T	*ABCA1*	0.357	-2.713	1.20×10^{-11}	0.168	0.507	0.741	0.190	0.476	0.689
rs662799	11q23.3	G/A	*APOA5*	0.649	5.713	1.12×10^{-46}	0.188	0.498	0.706	1.377	0.480	0.004
rs651821	11q23.3	C/T	*APOA5*	0.648	5.684	1.34×10^{-46}	0.187	0.496	0.706	1.350	0.478	0.005

rs11216126	11q23.3	A/C	*LOC101929011, BUD13*	0.157	4.271	1.93×10^{-16}	-0.365	0.640	0.569	1.032	0.637	0.106
rs10790162	11q23.3	A/G	*BUD13*	0.721	2.901	9.62×10^{-12}	0.144	0.537	0.789	1.037	0.512	0.043
rs964184	11q23.3	G/C	*ZPR1*	0.718	2.869	1.57×10^{-11}	0.140	0.536	0.794	0.990	0.511	0.053
rs6589566	11q23.3	G/A	*ZPR1*	0.721	2.872	1.57×10^{-11}	0.160	0.538	0.766	0.989	0.512	0.053
rs2266788	11q23.3	G/A	*APOA5*	0.721	2.869	1.67×10^{-11}	0.089	0.538	0.869	1.017	0.513	0.047
rs2075290	11q23.3	C/T	*ZPR1*	0.717	2.74	1.09×10^{-10}	0.250	0.536	0.641	0.911	0.512	0.075
rs2367970	11q23.3	G/A	*LOC101929011, BUD13*	0.351	-2.393	2.97×10^{-9}	-0.467	0.502	0.351	-0.975	0.481	0.043
rs1800588	15q21.3	C/T	*LIPC*	0.510	4.447	1.76×10^{-31}	0.482	0.477	0.312	0.395	0.457	0.387
rs1077834	15q21.3	T/C	*LIPC*	0.516	4.437	4.64×10^{-31}	0.360	0.479	0.453	0.343	0.459	0.455
rs1077835	15q21.3	A/G	*LIPC*	0.517	4.433	6.98×10^{-31}	0.361	0.480	0.452	0.342	0.460	0.457
rs261334	15q21.3	G/C	*LIPC*	0.535	-4.171	1.39×10^{-27}	-0.139	0.481	0.773	0.062	0.460	0.893
rs261290	15q21.3	T/C	*AQP9, LIPC*	0.567	-3.663	1.41×10^{-21}	0.040	0.484	0.935	0.122	0.462	0.791
rs2043082	15q21.3	G/A	*AQP9, LIPC*	0.430	3.663	3.18×10^{-21}	-0.047	0.488	0.924	-0.155	0.466	0.739

rs77250403	15q21.3	GA/G	*AQP9, LIPC*	0.434	3.670	5.23×10^{-21}	-0.100	0.493	0.839	-0.171	0.469	0.715
rs261291	15q21.3	T/C	*AQP9, LIPC*	0.526	3.336	4.83×10^{-18}	-0.300	0.484	0.535	-0.422	0.464	0.362
rs1532085	15q21.3	A/G	*AQP9, LIPC*	0.403	-2.914	9.93×10^{-14}	0.022	0.486	0.964	0.576	0.472	0.222
rs10468017	15q21.3	C/T	*AQP9, LIPC*	0.215	3.209	5.10×10^{-12}	-0.591	0.592	0.319	0.098	0.564	0.862
rs4775041	15q21.3	G/C	*AQP9, LIPC*	0.214	3.149	1.30×10^{-11}	-0.599	0.593	0.312	0.102	0.564	0.857
rs8034802	15q21.3	T/A	*LIPC*	0.666	2.552	4.40×10^{-10}	-0.027	0.510	0.957	-0.666	0.491	0.175
rs16940212	15q21.3	G/T	*AQP9, LIPC*	0.316	2.262	3.91×10^{-8}	-0.291	0.522	0.577	0.438	0.501	0.382
rs12148780	15q21.3	A/G	*AQP9, LIPC*	0.153	-2.902	3.95×10^{-8}	1.003	0.665	0.131	0.772	0.640	0.227
rs3764261	16q13	C/A	*HERPUD1, CETP*	0.207	8.863	6.07×10^{-82}	0.817	0.576	0.156	1.357	0.558	0.015
rs183130	16q13	C/T	*HERPUD1, CETP*	0.207	8.858	9.92×10^{-82}	0.783	0.576	0.174	1.338	0.558	0.017
rs247617	16q13	C/A	*HERPUD1, CETP*	0.207	8.848	1.86×10^{-81}	0.791	0.577	0.170	1.336	0.559	0.017
rs17231506	16q13	C/T	*HERPUD1, CETP*	0.209	8.847	2.14×10^{-81}	0.833	0.576	0.148	1.324	0.559	0.018

rs821840	16q13	A/G	*HERPUD1, CETP*	0.207	8.843	4.71×10^{-81}	0.791	0.577	0.171	1.318	0.560	0.019
rs247616	16q13	C/T	*HERPUD1, CETP*	0.207	8.812	6.28×10^{-81}	0.790	0.576	0.170	1.336	0.558	0.017
rs72786786	16q13	G/A	*HERPUD1, CETP*	0.206	8.409	6.40×10^{-58}	0.996	0.655	0.128	1.024	0.633	0.106
rs1532624	16q13	C/A	*CETP*	0.310	6.188	4.31×10^{-52}	0.973	0.512	0.057	0.674	0.487	0.167
rs711752	16q13	G/A	*CETP*	0.401	5.047	2.54×10^{-38}	1.226	0.489	0.012	1.238	0.471	0.009
rs173539	16q13	C/T	*HERPUD1, CETP*	0.316	4.970	4.84×10^{-34}	0.798	0.515	0.122	1.284	0.494	0.009
rs1864163	16q13	G/A	*CETP*	0.102	-7.442	2.75×10^{-32}	0.458	0.785	0.560	-0.922	0.726	0.204
rs7499892	16q13	C/T	*CETP*	0.164	-5.43	1.98×10^{-26}	0.077	0.640	0.904	-0.436	0.608	0.473
rs12708980	16q13	T/G	*CETP*	0.071	-6.418	1.25×10^{-17}	0.478	0.906	0.598	-0.833	0.863	0.335
rs1800775	16q13	C/A	*CETP*	0.552	3.162	1.61×10^{-16}	0.951	0.477	0.046	1.497	0.457	0.001
rs9989419	16q13	A/G	*HERPUD1, CETP*	0.745	2.871	2.81×10^{-8}	0.434	0.657	0.509	0.956	0.636	0.133
rs3786247	18q21.1	T/G	*LIPG*	0.460	3.209	1.02×10^{-16}	0.209	0.490	0.670	0.265	0.465	0.569

rs12970066	18q21.1	C/G	*LIPG*	0.253	2.748	6.04×10^{-10}	0.505	0.557	0.364	-0.093	0.533	0.862
rs35816125	18q21.1	C/G	*LIPG*	0.250	2.645	3.99×10^{-9}	0.334	0.558	0.549	0.172	0.541	0.750
rs429358	19q13.32	T/C	*APOE*	0.100	-3.594	2.98×10^{-8}	0.069	0.802	0.931	-0.276	0.779	0.723

ALT, alternative allele; BMI, body mass index; Coeff., coefficient; GWAS, genome-wide association study; HDL-C, high density lipoprotein cholesterol; REF, referent allele; SE, standard error; SNP, single nucleotide polymorphism.

[a] Adjusted for age, sex, smoking, drinking, daily activity, habitual exercise, egg intake, and BMI. The coefficient value represents change in HDL-C per ALT allele copy (0, 1, 2) for the SNP.

VITA

Yora Nindita, was born in Surakarta, Indonesia. In 2000, she entered the Medical Faculty of Diponegoro University, Semarang and received her Bachelor of Medicine degree in 2004. She had her clinical experience in RSUP Dr. Kariadi and received her medical doctor degree in 2006. She started working in the Department of Pharmacology and Therapeutic, Faculty of Medicine, Diponegoro University in January 2008. She received the degree of Master of Science majoring Pharmacology and Toxicology from the Graduate School of Basic Medical Science and Biomedical Gadjah Mada University in 2011. In October 2013, she entered the Graduate School of Medical and Dental Sciences, Kagoshima University. While taking the Master (October 2009 - October 2011) and Ph.D degree (October 2013 – September 2017), she takes study assignment leaves from her office. She re-started working in the Department of Pharmacology and Therapeutic, Faculty of Medicine, Diponegoro University in January 2018.

Address: nindita.yora@gmail.com