Genomics 111 (2019) 520–533

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Original Article

Identification of six novel susceptibility loci for dyslipidemia using longitudinal exome-wide association studies in a Japanese population

Yoshiki Yasukochi[a,b,*], Jun Sakuma[b,c,d], Ichiro Takeuchi[b,d,e], Kimihiko Kato[a,f], Mitsutoshi Oguri[a,g], Tetsuo Fujimaki[h], Hideki Horibe[i], Yoshiji Yamada[a,b]

[a] Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu 514-8507, Japan
[b] CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan
[c] Computer Science Department, College of Information Science, University of Tsukuba, Tsukuba 305-8573, Japan
[d] RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan
[e] Department of Computer Science, Nagoya Institute of Technology, Nagoya 466-8555, Japan
[f] Department of Internal Medicine, Meitoh Hospital, Nagoya 465-0025, Japan
[g] Department of Cardiology, Kasugai Municipal Hospital, Kasugai 486-8510, Japan
[h] Department of Cardiovascular Medicine, Inabe General Hospital, Inabe 511-0428, Japan
[i] Department of Cardiovascular Medicine, Gifu Prefectural Tajimi Hospital, Tajimi 507-8522, Japan

ARTICLE INFO

Keywords:
Dyslipidemia
Exome-wide association study
Generalized estimating equation
LDL-cholesterol
HDL-cholesterol
Triglyceride

ABSTRACT

Recent genome-wide association studies have identified various dyslipidemia-related genetic variants. However, most studies were conducted in a cross-sectional manner. We thus performed longitudinal exome-wide association studies of dyslipidemia in a Japanese population. We used ~244,000 genetic variants and clinical data of 6022 Japanese individuals who had undergone annual health checkups for several years. After quality control, the association of dyslipidemia-related phenotypes with 24,691 single nucleotide polymorphisms (SNPs) was tested using the generalized estimating equation model. In total, 82 SNPs were significantly ($P < 2.03 \times 10^{-6}$) associated with dyslipidemia phenotypes. Of these SNPs, four (rs74416240 of *TCHP*, rs925368 of *GIT2*, rs7969300 of *ATXN2*, and rs12231744 of *NAA25*) and two (rs34902660 of *SLC17A3* and rs1042127 of *CDSN*) were identified as novel genetic determinants of hypo-HDL- and hyper-LDL-cholesterolemia, respectively. A replication study using the cross-sectional data of 8310 Japanese individuals showed the association of the six identified SNPs with dyslipidemia-related traits.

1. Introduction

Dyslipidemia is defined as the abnormal serum concentrations of triglycerides, low-density lipoprotein (LDL)-cholesterol, and high-density lipoprotein (HDL)-cholesterol; hypertriglyceridemia and hyper-LDL-cholesterolemia indicate the increased serum concentrations of triglycerides and LDL-cholesterol, respectively, while hypo-HDL-cholesterolemia indicates a decreased serum concentration of HDL-cholesterol. The National Health and Nutrition Examination Survey data has reported a high prevalence of dyslipidemia (53%) in 2003–2006 among adults in the United States [1]. Meanwhile the Ministry of Health, Labour and Welfare has reported that 21.8% of the Japanese adults were suspected to have dyslipidemia in 2016 [2]. Dyslipidemia can be a risk factor for various disorders, including coronary heart disease and ischemic stroke [3–5]. Therefore, the investigation of susceptibility genes for dyslipidemia is crucial for personalized disease prevention.

Previous studies have identified a number of genetic variants affecting the serum concentrations of triglycerides, LDL-cholesterol, and HDL-cholesterol. A large-scale meta-analysis of genome-wide association studies (GWASs) of plasma lipids has identified 95 loci related to lipid traits, and 59 of the identified loci were newly reported in the study [6]. Another large-scale meta-analysis conducted to further assess the genetic association of these loci with plasma lipid phenotypes, has identified candidate genetic variants in 21 previously unreported genes [7]. Furthermore, a two-stage association study for LDL-cholesterol focusing upon rare or low-frequency variants in protein-coding regions has revealed significant relations between LDL-cholesterol and additional unreported genetic variants [8]. Although previous studies have identified common genetic variants that confer susceptibility to

* Corresponding author at: Department of Human Functional Genomics, Advanced Science Research Promotion Center, Organization for the Promotion of Regional Innovation, Mie University, 1577 Kurima-machiya, Tsu, Mie 514-8507, Japan.
E-mail address: hyasukou@proof.ocn.ne.jp (Y. Yasukochi).

https://doi.org/10.1016/j.ygeno.2018.05.015
Received 11 March 2018; Received in revised form 9 May 2018; Accepted 18 May 2018
Available online 04 June 2018

dyslipidemia across various ethnic groups, the interethnic differences in the disease susceptibility loci have been observed [9, 10]. Therefore, it is possible that novel susceptibility loci for dyslipidemia remain to be identified definitively in Japanese.

Conventional GWASs have been conducted using a cross-sectional study design that commonly measures traits at a single point in time. Therefore, subjects who are likely to be affected by dyslipidemia in the future may be regarded as controls at the time of data collection. To address this issue, we traced the blood lipid profiles of 6022 Japanese individuals who had undergone annual health checkups over several years. To identify dyslipidemia-related genetic variants, we performed longitudinal exome-wide association studies of the prevalence of hypertriglyceridemia, hyper-LDL-cholesterolemia, or hypo-HDL-cholesterolemia as well as the serum concentrations of triglycerides, LDL-cholesterol, or HDL-cholesterol.

2. Results

2.1. Characteristics of subjects in the discovery cohort

We measured the blood lipid profiles of 6022 community-dwelling individuals who visited the Health Care Center at Inabe General Hospital, Inabe City, Japan, for an annual health checkup from April 2003 to March 2014 (mean follow-up period, 5 years). We refer to this cohort as "discovery cohort." The characteristics of the 6022 subjects in the discovery cohort concerning the longitudinal data are summarized in Table 1. The prevalence of dyslipidemia based on longitudinal data was higher in males than in females; the proportion of males was 73.5%, 83.9%, and 55.6% among subjects with hypertriglyceridemia, hypo-HDL-cholesterolemia, and hyper-LDL-cholesterolemia, respectively. Compared with the controls, the prevalence of other complex disorders (hypertension, type 2 diabetes mellitus, chronic kidney disease, and hyperuricemia) was higher in subjects with dyslipidemia. Most physiological or clinical parameters examined [age, weight, body mass index (BMI), waist circumference, systolic and diastolic blood pressure, fasting plasma glucose level, blood hemoglobin A_{1c} content, and serum concentrations of creatinine and uric acid] were higher, whereas estimated glomerular filtration rate was lower in the subjects with dyslipidemia than in the controls.

2.2. Longitudinal exome-wide association study for hypertriglyceridemia

In our longitudinal exome-wide association studies of the 6022 subjects in the discovery cohort, the relations of 24,691 single nucleotide polymorphisms (SNPs) were tested for three inheritance models with six dyslipidemia-related parameters (hypertriglyceridemia; hypo-HDL-cholesterolemia; hyper-LDL-cholesterolemia; and serum concentrations of triglycerides, HDL-cholesterol, or LDL-cholesterol) using the generalized estimating equation (GEE) models after adjusting for age, sex, BMI, and smoking status. SNPs with statistical significance ($P < 2.03 \times 10^{-6}$) were further filtered out by the threshold of approxdf (a scale of small effective sample size) of > 30. The study design and results of the longitudinal exome-wide association studies are shown in Fig. 1 and Fig. 2 or Supplementary Tables S1–S3, respectively.

In the longitudinal exome-wide association studies for hypertriglyceridemia, 26 different SNPs were significantly ($P < 2.03 \times 10^{-6}$, approxdf > 30) associated with the prevalence of hypertriglyceridemia in the three inheritance models, whereas another set of 26 SNPs was significantly associated with the serum triglyceride concentration (Fig. 2 and Supplementary Tables S1–S3). Between these two sets of 26 SNPs, 22 were overlapped. Among 30 SNPs identified, 28 had been previously reported to be associated with dyslipidemia-related phenotypes according to Genome-Wide Repository of Associations Between SNPs and Phenotypes (GRASP), GWAS Catalogue, and DisGeNET databases. The remaining two SNPs were in a linkage disequilibrium (LD) with SNPs that have been reported to be associated with dyslipidemia phenotypes ($r^2 > 0.85$), according to LDproxy analysis, an LDlink web-based application, for JPT (Japanese in Tokyo, Japan) from the 1000 Genomes Project. Hence, these two SNPs were not considered as novel susceptibility loci for hypertriglyceridemia.

2.3. Longitudinal exome-wide association study for hypo-HDL-cholesterolemia

The GEE models with adjustments for age, sex, BMI, and smoking status revealed 21 SNPs significantly ($P < 2.03 \times 10^{-6}$, approxdf > 30) associated with the prevalence of hypo-HDL-cholesterolemia in the dominant and additive models (Fig. 2 and Supplementary Tables S1–S3). In the longitudinal exome-wide association studies of serum HDL-cholesterol concentration, 45 SNPs were significantly associated in the three inheritance models. Between the two sets of 21 and 45 SNPs, 20 were overlapped. Thus, 46 SNPs were associated with HDL-cholesterol-related phenotypes in the discovery cohort.

Of the 46 SNPs, 8 had not been identified as significantly associated with dyslipidemia-related phenotypes according to the three databases. LDs between the eight SNPs and previously identified SNPs in JPT from the 1000 Genomes Project were surveyed using LDproxy and GRASP database (see Materials and methods). An r^2 value was adopted as a marker to estimate the strength of LD because r^2 is directly linked to statistical power for detecting an association between the prevalence of disease and the allele frequencies of the target and adjacent SNPs [11]. On the basis of the previous studies [11, 12], an r^2 of > 0.5 was considered to be a significant LD. The analysis using LDproxy and GRASP indicated that four of the eight SNPs were in LD ($r^2 > 0.5$) with SNPs previously identified associated with dyslipidemia. Thus, these SNPs were not considered as novel susceptibility loci. Consequently, the remaining four SNPs [rs74416240 of *TCHP* ($P = 6.2 \times 10^{-9}$, approxdf = 76), rs925368 of *GIT2* ($P = 6.4 \times 10^{-9}$, approxdf = 69), rs7969300 of *ATXN2* ($P = 7.0 \times 10^{-8}$, approxdf = 266–466), and rs12231744 of *NAA25* ($P = 3.1 \times 10^{-7}$, approxdf = 283–369)] were identified as novel genetic determinants of hypo-HDL-cholesterolemia. These SNPs are closely located in the chromosomal region 12q24.11–q24.13 (Fig. 3).

In the present study, LD was observed between rs74416240 and rs925368 ($r^2 = 0.94$) as well as between rs7969300 and rs12231744 ($r^2 = 0.68$) (Fig. 3), suggesting that the effects of these SNPs on serum HDL-cholesterol concentration were not independent. Furthermore, r^2 values between the four identified and some adjacent SNPs were ~0.3, indicating a weak LD among these SNPs. Thus, the association of the four identified SNPs with dyslipidemia might be affected to a certain extent by other adjacent SNPs that confer susceptibility to dyslipidemia. The association between the four identified SNPs and hypo-HDL-cholesterolemia-related phenotypes should be carefully interpreted, and further functional analyses are required to verify the association.

2.4. Longitudinal exome-wide association study for hyper-LDL-cholesterolemia

Our longitudinal exome-wide association studies revealed that two SNPs were significantly ($P < 2.03 \times 10^{-6}$, approxdf > 30) associated with the prevalence of hyper-LDL-cholesterolemia in the dominant and additive models. Furthermore, 24 SNPs were significantly associated with serum LDL-cholesterol concentration (Fig. 2 and Supplementary Tables S1–S3). Of the 26 associated SNPs, rs34902660 of *SLC17A3* was associated with both the prevalence of hyper-LDL-cholesterolemia and the serum concentration of LDL-cholesterol. In total, 25 SNPs were significantly associated with phenotypes related to LDL-cholesterol. According to the analysis using LDproxy and GRASP database (see above), two [rs34902660 of *SLC17A3* ($P = 1.8 \times 10^{-7}$ to 1.2×10^{-6}, approxdf = 31–34) and rs1042127 of *CDSN* ($P = 2.4 \times 10^{-7}$ to 1.3×10^{-6}, approxdf = 162)] of the 25 SNPs were identified as novel susceptibility loci for hyper-LDL-cholesterolemia.

Table 1
Longitudinal characteristics of study subjects in the discovery.

Characteristic	Controls[a]	Hyper-triglyceridemia[a]	Controls[a]	Hypo-HDL-cholesterolemia[a]	Controls[a]	Hyper-LDL-cholesterolemia[a]
No. of subjects	3966[b]	2056[b]	5324[b]	698[b]	3251[b]	2769[b]
Male (%)	47.9 (9398)[c]	73.5 (6545)[c]	52.2 (13,169)[c]	83.9 (2774)[c]	56.1 (9619)[c]	55.6 (6314)[c]
Female (%)	52.1 (10,217)[c]	26.5 (2360)[c]	47.8 (12,046)[c]	16.1 (531)[c]	43.9 (7537)[c]	44.4 (5040)[c]
Age (years)	51.6 ± 12.5 (19,615)	54.5 ± 11.3 (8905)	52.2 ± 12.2 (25,215)	54.7 ± 12.5 (3305)	51.0 ± 12.8 (17,156)	54.7 ± 10.9 (11,354)
Height (cm)	161.6 ± 8.9 (19,213)	164.7 ± 9.1 (8699)	162.0 ± 9.0 (24,812)	166.7 ± 8.4 (3100)	162.9 ± 8.9 (16,666)	162.0 ± 9.3 (11,244)
Weight (kg)	58.5 ± 10.8 (19,211)	66.7 ± 12.4 (8699)	59.9 ± 11.4 (24,811)	69.7 ± 12.7 (3099)	60.1 ± 11.8 (16,664)	62.4 ± 12.1 (11,244)
Body mass index (kg/m^2)	22.3 ± 3.1 (19,211)	24.5 ± 3.4 (8699)	22.7 ± 3.2 (24,811)	25.0 ± 3.4 (3099)	22.5 ± 3.3 (16,664)	23.6 ± 3.3 (11,244)
Waist circumference (cm)	78.6 ± 8.7 (14,467)	85.2 ± 8.6 (6880)	80.0 ± 9.1 (18,925)	86.5 ± 8.4 (2422)	79.4 ± 9.3 (12,132)	82.5 ± 8.8 (9213)
Smoking (%)	33.7 (19,615)	54.4 (8905)	37.3 (25,215)	62.2 (3305)	40.6 (17,156)	39.5 (11,354)
Hypertension (%)	27.5 (19,615)	47.6 (8905)	32.0 (25,215)	47.5 (3305)	31.0 (17,156)	37.9 (11,354)
Systolic blood pressure (mm Hg)	118.7 ± 16.2 (19,206)	125.1 ± 15.9 (8694)	118.8 ± 16.4 (24,804)	123.1 ± 15.9 (3096)	119.7 ± 16.5 (16,659)	122.4 ± 16.0 (11,240)
Diastolic blood pressure (mm Hg)	72.9 ± 12.1 (19,206)	78.6 ± 11.8 (8694)	73.1 ± 12.3 (24,804)	77.1 ± 12.0 (3096)	74.0 ± 12.4 (16,659)	75.9 ± 12.0 (11,240)
Type 2 diabetes mellitus (%)	8.9 (19,615)	19.9 (8905)	10.5 (25,215)	26.2 (3305)	10.9 (17,156)	14.5 (11,354)
Fasting plasma glucose (mmol/L)	5.48 ± 0.95 (19,200)	5.94 ± 1.45 (8869)	5.49 ± 1.08 (24,787)	6.05 ± 1.52 (3282)	5.55 ± 1.08 (16,737)	5.72 ± 1.24 (11,322)
Blood hemoglobin A_{1c} (%)	5.64 ± 0.56 (14,503)	5.88 ± 0.81 (6504)	5.64 ± 0.63 (18,685)	5.94 ± 0.79 (2322)	5.65 ± 0.62 (12,223)	5.82 ± 0.69 (8782)
Serum triglycerides (mmol/L)	0.89 ± 0.32 (19,172)	2.06 ± 1.21 (8857)	0.93 ± 0.74 (24,767)	2.04 ± 1.49 (3262)	1.13 ± 0.94 (16,705)	1.42 ± 0.84 (11,317)
Serum HDL-cholesterol (mmol/L)	1.72 ± 0.42 (19,157)	1.37 ± 0.35 (8837)	1.73 ± 0.40 (24,737)	1.06 ± 0.21 (3257)	1.65 ± 0.45 (16,675)	1.55 ± 0.39 (11,312)
Serum LDL-cholesterol (mmol/L)	3.09 ± 0.76 (18,389)	3.39 ± 0.86 (8433)	3.11 ± 0.80 (23,732)	3.20 ± 0.84 (3090)	2.88 ± 0.53 (15,515)	3.79 ± 0.70 (11,300)
Chronic kidney disease (%)	9.0 (19,615)	15.6 (8905)	9.8 (25,215)	21.0 (3305)	10.3 (17,156)	12.1 (11,354)
Serum creatinine (μmol/L)	74.5 ± 100.7 (17,452)	83.1 ± 109.4 (8307)	70.7 ± 79.6 (22,708)	121.2 ± 203.5 (3051)	79.9 ± 126.1 (15,108)	69.6 ± 55.0 (10,641)
eGFR ($mL\,min^{-1}\,1.73\,m^{-2}$)	79.7 ± 17.8 (17,452)	76.2 ± 17.7 (8307)	80.1 ± 16.5 (22,708)	72.6 ± 25.0 (3051)	79.3 ± 19.0 (15,108)	77.5 ± 16.0 (10,641)
Hyperuricemia (%)	12.8 (19,615)	28.7 (8905)	15.9 (25,215)	31.8 (3305)	17.6 (17,156)	17.9 (11,354)
Serum uric acid (μmol/L)	311.5 ± 82.2 (17,102)	366.2 ± 84.8 (8155)	312.1 ± 85.5 (22,256)	374.0 ± 83.6 (3001)	324.5 ± 88.9 (14,752)	334.1 ± 83.7 (10,495)

HDL, high density lipoprotein. LDL, low density lipoprotein. eGFR, estimated glomerular filtration rate.

[a] Quantitative data are means and standard deviations, and values in parentheses indicate the numbers of measurements taken.

[b] The numbers are based on data examined in the latest year.

[c] The proportion of males or females is calculated based on the numbers of measurements taken shown in parentheses.

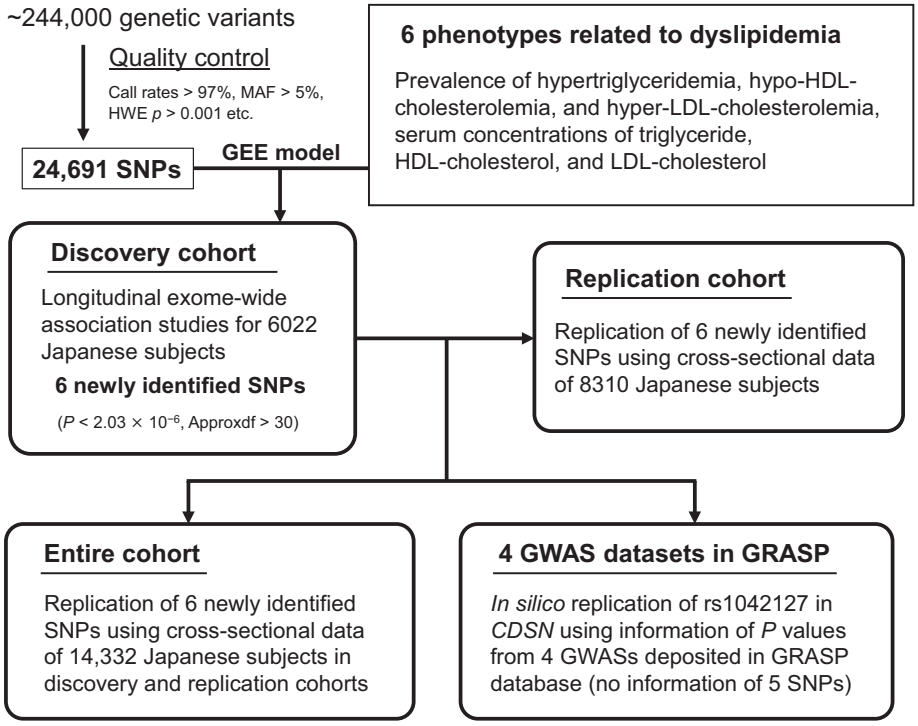

Fig. 1. Study flow chart. An approxdf is a scale of small effective sample size: approxdf = 2 × MAF × Nindep, where Nindep is the sum of the estimated number of independent observations per person. GEE, generalized estimating equation. GWAS, genome-wide association study. HDL, high density lipoprotein. HWE, Hardy–Weinberg equilibrium. LDL, low density lipoprotein. MAF, minor allele frequency. SNP, single nucleotide polymorphism.

In the discovery cohort, r^2 values between the two identified SNPs and some adjacent SNPs were ~0.4 (Fig. 4). The rs34902660 of *SLC17A3* showed a weak LD (r^2 = 0.41) with rs9267546 previously identified to be associated with serum triglyceride concentration. Thus, the association of rs34902660 with dyslipidemia might be affected to a certain extent by rs9267546. The association should be carefully

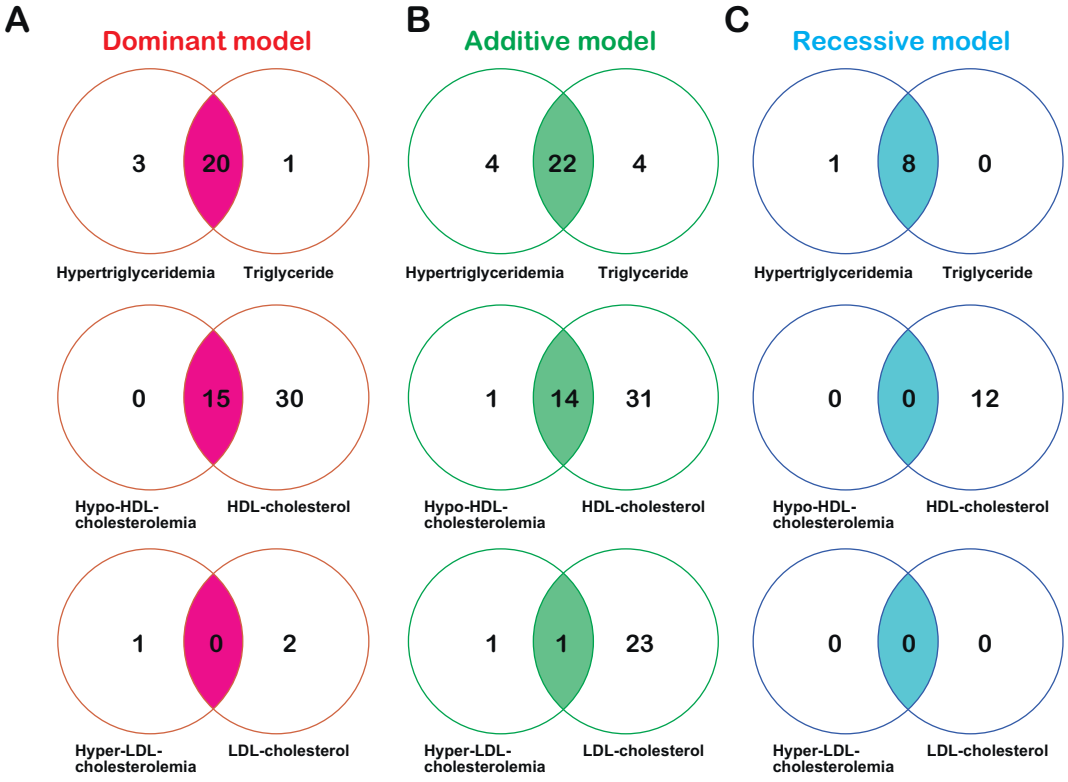

Fig. 2. SNPs showing significant association with dyslipidemia. The association was tested by the generalized estimating equation model with adjustments for age, sex, BMI, and smoking status in the (A) dominant, (B) additive, and (C) recessive inheritance models. The analyses independently tested the association between SNPs and six clinical parameters: prevalence of hypertriglyceridemia, hypo-HDL-cholesterolemia, and hyper-LDL-cholesterolemia or serum concentrations of triglycerides, HDL-cholesterol, and LDL-cholesterol. After employing Bonferroni's correction, $P < 2.03 \times 10^{-6}$ was considered statistically significant. HDL, high density lipoprotein. LDL, low density lipoprotein.

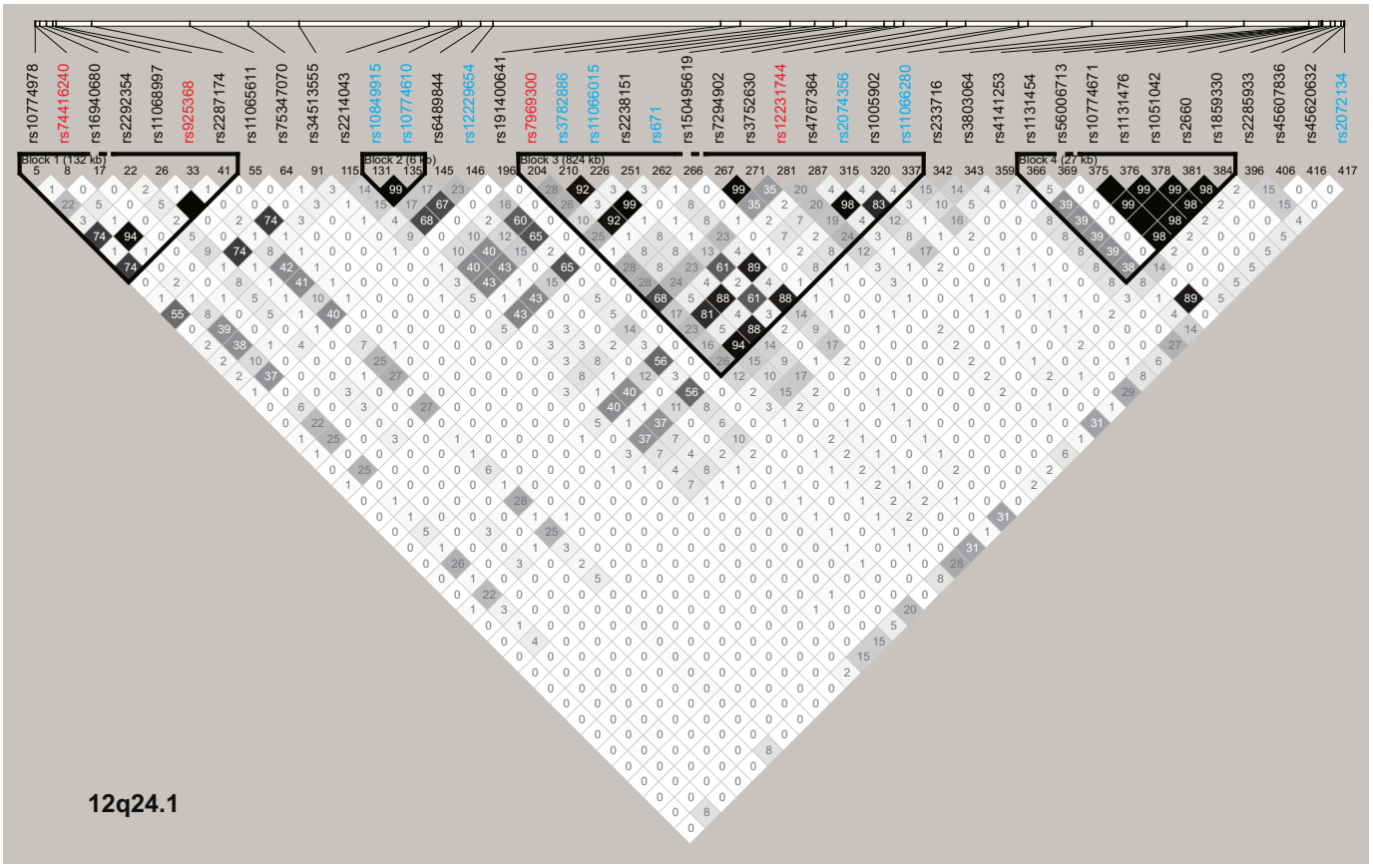

Fig. 3. Linkage disequilibrium map of 43 biallelic sites in a ~3.1 Mb genomic region at 12q24.11–q24.13 in the discovery cohort. The number in a diamond represents r^2 value (×100). The SNPs with a minor allele frequency of < 0.01 were excluded from the analysis. The dyslipidemia-related SNPs detected in the present and previous studies are indicated in blue. The novel dyslipidemia-related SNPs detected in the present study are indicated in red.

interpreted, and further functional analyses are warranted to verify the association.

2.5. Replication studies for candidate SNPs associated with dyslipidemia

The association of the six candidate SNPs identified in our longitudinal exome-wide association studies with the six dyslipidemia-related traits was examined using the cross-sectional data of 8310 Japanese subjects (see Materials and methods) in the replication cohort (Table 2). In this replication study, all six candidate SNPs were associated ($P < .05$) with at least three of the six phenotypes examined: the rs1042127 of *CDSN* was associated with the serum concentrations of triglycerides and HDL-cholesterol and the prevalence of hypo-HDL-cholesterolemia although this SNP was related to serum LDL-cholesterol concertation in the discovery cohort.

We also examined the association of the six candidate SNPs with the dyslipidemia-related traits using cross-sectional data of all Japanese subjects (14,332 subjects) in the combined cohort of discovery and replication cohorts (Table 2): we refer to this cohort as “entire cohort.” Here, in the discovery cohort, cross-sectional data of 6022 subjects in the latest year were used. In the entire cohort, the association of the candidate SNPs with dyslipidemia-related phenotypes showed good agreement with that observed in the discovery cohort: four (rs74416240, rs925368, rs7969300, and rs12231744) and two (rs34902660 and rs1042127) SNPs were significantly [$P < .008$ (0.05/6 SNPs)] associated with serum concentrations of HDL- and LDL-cholesterol, respectively.

Additionally, the association of the candidate SNPs detected in the discovery cohort with dyslipidemia phenotypes was examined using large datasets from four previous GWASs [13–16], data for which are deposited in GRASP database (Supplementary Table S4). Information regarding *P*-values was available for rs1042127 of *CDSN* but not for the other five SNPs because minor allele frequencies (MAFs) of these SNPs were < 0.3% in European populations on the basis of the allele frequency data of the 1000 Genomes Project (Table 3). In agreement with the results of our replication study, rs1042127 was not associated with serum LDL cholesterol concentration in the 4 GWAS datasets, but was associated with serum HDL cholesterol concentration in the studies of Teslovich et al. [13] (*P*-value = .032) and Willer et al. [16] (*P*-value = 3.5×10^{-4}) (Supplementary Table S4).

2.6. Comparisons of quantitative data among subjects with different genotypes of the six identified SNPs in the discovery cohort

Based on the blood lipid profiles of the subjects in the latest year, the quantitative values of dyslipidemia-related phenotypes were compared among different genotypes of the six identified SNPs in the discovery cohort (Table 4). In comparisons of the four candidate SNPs associated with hypo-HDL-cholesterolemia, serum HDL-cholesterol concentration was significantly ($P < .008$, by one-way analysis of variance) lower in the major alleles [A (rs925368), T (rs7969300), and C (rs12231744)] than in the corresponding minor alleles (G, C, and T). In contrast, serum HDL-cholesterol concentration was significantly lower in the minor allele (A) of rs74416240 than in the major allele (G). These results suggest that the major alleles of rs925368, rs7969300, and rs12231744 and the minor allele of rs74416240 may be risk factors for hypo-HDL-cholesterolemia. Serum LDL-cholesterol concentrations were significantly higher in subjects with the minor alleles of two candidate SNPs [A (rs34902660) and G (rs1042127)] associated with hyper-LDL-cholesterolemia than in the corresponding major alleles (C

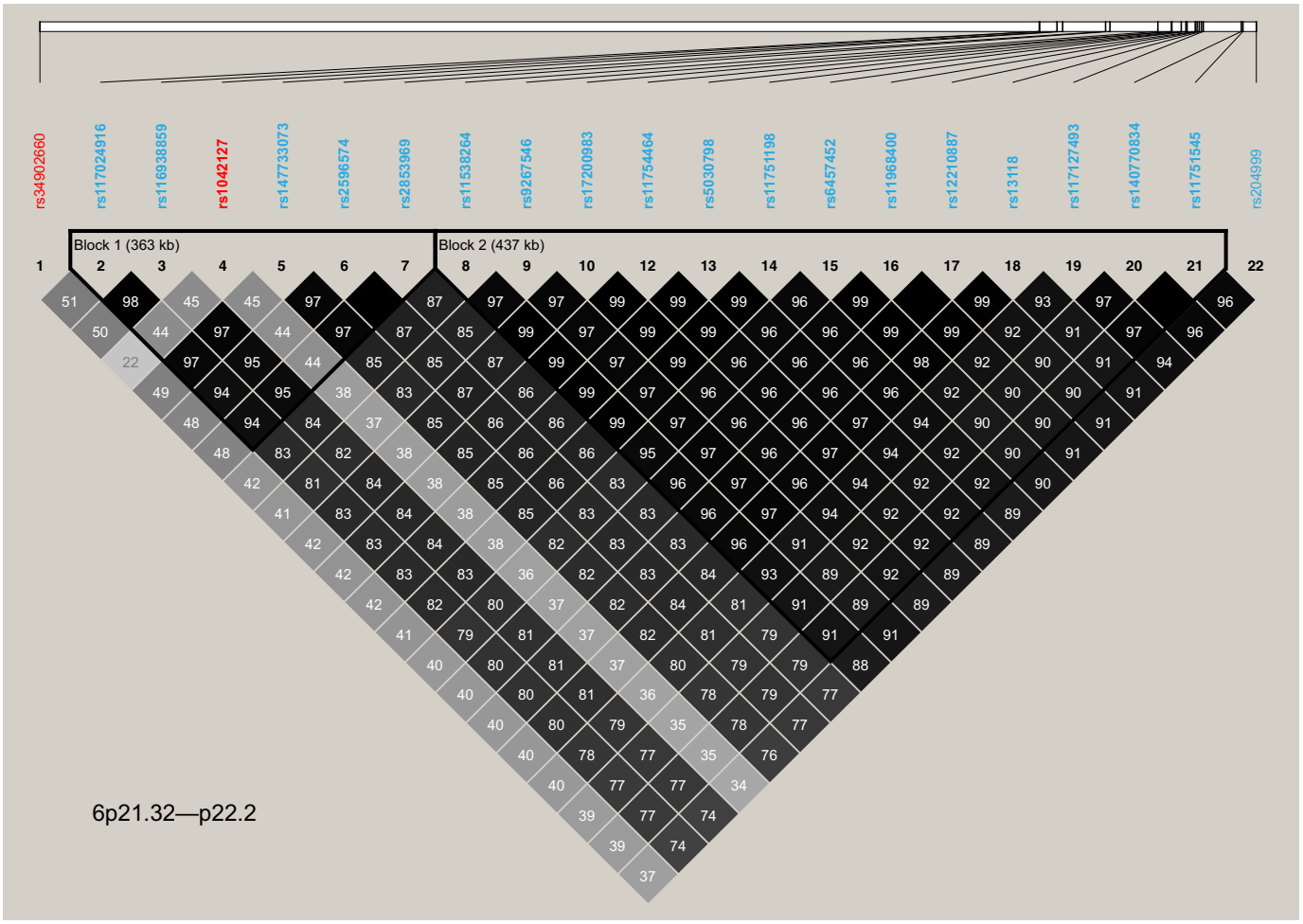

Fig. 4. Linkage disequilibrium map of 21 biallelic sites in a ~6.2 Mb genomic region at 6p21.32–p22.2 in the discovery cohort. The number in a diamond represents r^2 value (×100). The SNPs with a minor allele frequency of < 0.01 were excluded from the analysis. The dyslipidemia-associated SNPs detected in the present and previous studies are shown in blue. The novel dyslipidemia-related SNPs detected in the present study are shown in red.

and T), suggesting that these minor alleles may be risk factors for hyper-LDL-cholesterolemia.

We examined the relation of each candidate SNP to serum concentrations of HDL-cholesterol or LDL-cholesterol in the replication cohort (Table 5). The serum concentrations of HDL-cholesterol were significantly ($P < .008$) different among genotypes in the four SNPs associated with hypo-HDL-cholesterolemia, whereas the two SNPs (rs34902660 and rs1042127) associated with hyper-LDL-cholesterolemia were not significantly related to the serum LDL-cholesterol levels. This discrepancy may be attributable to the effect of medical treatment, given that subjects in the replication study were recruited from individuals either who visited outpatient clinics of or were admitted to participating hospitals. We have thus examined the relation of six SNPs to the serum levels of HDL-cholesterol or LDL-cholesterol in the entire cohort. The serum LDL- or HDL-cholesterol levels were significantly different among genotypes in all the six SNPs (Table 5). The relations of six SNPs to serum LDL- or HDL-cholesterol levels in discovery and replication cohorts have been shown as box-plots in Figs. 5 and 6, respectively.

3. Discussion

In our longitudinal exome-wide association studies for dyslipidemia-related traits, 82 SNPs were significantly associated with at least one of the six phenotypes. Of the 82 SNPs, four (rs74416240 of *TCHP*, rs925368 of *GIT2*, rs7969300 of *ATXN2*, and rs12231744 of *NAA25*) and two (rs34902660 of *SLC17A3* and rs1042127 of *CDSN*) were identified as novel susceptibility loci for hypo-HDL-cholesterolemia and hyper-LDL-cholesterolemia, respectively. The association of these SNPs with dyslipidemia-related phenotypes was replicated using the cross-sectional data of 8310 Japanese individuals as well as 14,332 individuals in the entire cohort. The discovery and replication cohorts are comprised of only Japanese individuals, and five of six novel susceptibility loci identified in the present study were monomorphic or rare variants in European populations. Therefore, our findings may not be generalizable to other racial or ethnic groups such as European Americans.

In the discovery cohort, rs74416240 of *TCHP* and rs925368 of *GIT2* were in LD ($r^2 = 0.94$). These SNPs were significantly associated with serum HDL-cholesterol concentration. A previous GWAS has shown the association ($P = 7.0 \times 10^{-6}$) between another SNP, rs2292354 of *GIT2*, and serum HDL-cholesterol concentrations in Asian Indian males [17]. Thus, *GIT2* might be more important for lipid metabolisms than *TCHP* although it is difficult to distinguish which SNP is responsible for the lipid metabolisms. *GIT2* is a member of G-protein-coupled receptor-kinase-interacting proteins involved in diverse cellular processes, such as the regulation of membrane trafficking between the plasma membrane and endosomes [18]. According to The Human Protein Atlas (http://www.proteinatlas.org/), *GIT2* is ubiquitously expressed in various organs and tissues. Given that molecules related to lipid metabolism are involved in membrane transport processes [19], *GIT2* might affect alterations in HDL-cholesterol concentration although further

Table 2
Association of candidate SNPs detected in the discovery cohort with cross-sectional data of dyslipidemia-related phenotypes in ~8310 Japanese subjects in the replication cohort and in ~14,332 Japanese subjects in the entire cohort as determined by the chi-square test (for categorical data) or linear regression analysis (for quantitative data).

RefSNP ID	Gene	Model	Replication[a]	Entire[b]	Replication[a]	Entire[b]	Replication[a]	Entire[b]	Replication[a]	Entire[b]	Replication[a]	Entire[b]	Replication[a]	Entire[b]
			Hyper-triglyceridemia		Triglyceride		Hypo-HDL-cholesterolemia		HDL-cholesterol		Hyper-LDL-cholesterolemia		LDL-cholesterol	
			P-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value	*P*-value
rs34902660	*SLC17A3*	Additive	**0.015**	**1.65×10^{-4}**	**1.25×10^{-5}**	**9.4×10^{-5}**	**0.011**	0.115	**3.08×10^{-5}**	0.464	0.134	**1.19×10^{-8}**	**0.033**	**1.99×10^{-5}**
		Dominant	**0.012**	**3.22×10^{-5}**	**3.06×10^{-5}**	**2.1×10^{-4}**	**0.023**	**0.040**	**5.82×10^{-5}**	0.216	0.136	**2.07×10^{-9}**	**0.045**	**8.75×10^{-6}**
		Recessive	1.000	0.609	**0.028**	**0.038**	0.174	0.940	0.056	0.753	0.871	**0.047**	0.238	**0.024**
rs1042127	*CDSN*	Additive	0.684	0.461	**0.003**	0.138	0.079	0.398	0.108	0.129	0.263	**8.60×10^{-5}**	0.163	**0.001**
		Dominant	0.474	0.462	**0.004**	0.073	0.066	0.317	**0.014**	0.731	0.373	**1.84×10^{-5}**	0.156	**1.88×10^{-4}**
		Recessive	0.487	0.454	0.087	0.180	0.950	0.556	0.198	0.070	0.464	0.069	0.560	0.071
rs74416240	*TCHP*	Additive	0.093	0.298	0.363	0.939	**0.013**	**0.026**	**1.40×10^{-6}**	**2.22×10^{-5}**	**0.015**	**5.95×10^{-4}**	0.341	0.186
		Dominant	0.067	0.542	0.278	0.834	**0.039**	**0.018**	**5.89×10^{-6}**	**9.11×10^{-5}**	**0.011**	**1.21×10^{-4}**	0.319	0.069
		Recessive	0.700	0.212	0.776	0.817	0.094	0.068	**0.007**	**9.43×10^{-4}**	0.913	0.526	0.863	0.517
rs925368	*GIT2*	Additive	0.071	0.182	0.315	0.560	**0.011**	**0.031**	**3.85×10^{-6}**	**1.26×10^{-4}**	**0.026**	**4.49×10^{-4}**	0.355	0.150
		Dominant	**0.050**	0.425	0.208	0.662	**0.030**	**0.013**	**8.15×10^{-6}**	**8.15×10^{-5}**	**0.017**	**8.86×10^{-5}**	0.267	0.053
		Recessive	0.753	0.150	0.557	0.390	0.112	0.156	**0.034**	**0.016**	0.721	0.506	0.712	0.811
rs7969300	*ATXN2*	Additive	0.726	**0.023**	0.866	0.409	**0.001**	**1.98×10^{-4}**	**3.20×10^{-8}**	**1.98×10^{-9}**	**0.001**	0.055	**0.002**	**0.026**
		Dominant	0.839	0.738	0.977	0.519	**0.039**	**4.19×10^{-4}**	**4.71×10^{-6}**	**2.86×10^{-7}**	**0.009**	**0.017**	**0.001**	**0.007**
		Recessive	0.300	**0.007**	0.776	0.187	**0.007**	**0.001**	**7.98×10^{-6}**	**2.25×10^{-7}**	0.066	0.272	0.136	0.324
rs12231744	*NAA25*	Additive	0.815	0.203	0.428	0.887	**2.25×10^{-4}**	**6.03×10^{-5}**	**2.10×10^{-7}**	**2.42×10^{-10}**	**0.002**	0.108	**0.002**	0.062
		Dominant	0.416	0.653	0.391	0.755	**0.015**	**6.27×10^{-4}**	**4.47×10^{-6}**	**1.55×10^{-7}**	**0.041**	0.214	**0.001**	**0.037**
		Recessive	0.284	0.075	0.734	0.798	**0.006**	**1.86×10^{-4}**	**1.53×10^{-4}**	**2.52×10^{-8}**	**0.017**	**0.044**	0.135	0.087

P-values of < 0.05 are shown in bold. HDL, high density lipoprotein. LDL, low density lipoprotein.
[a] Replication cohort. [b] Combined cohort of discovery and replication cohorts.

Table 3
Allele frequencies of six dyslipidemia-associated SNPs in human populations.

Chr.	RefSNP ID	East Asian							South Asian[a]	European[a]	African[a]
	Gene	All	JP-Inabe	JPT[a]	CDX[a]	CHB[a]	CHS[a]	KHV[a]			
Chr. 6	rs34902660	C: 0.907 (11,831)	C: 0.900 (10,836)	C: 0.938 (195)	C: 1.000 (186)	C: 1.000 (206)	C: 1.000 (210)	C: 1.000 (198)	C: 1.000 (978)	C: 1.000 (1006)	C: 0.984 (1301)
	SLC17A3	A: 0.093 (1219)	A: 0.100 (1206)	A: 0.062 (13)							A: 0.016 (21)
	rs1042127	T: 0.785 (12,841)	T: 0.777 (9362)	T: 0.822 (171)	T: 0.785 (146)	T: 0.840 (173)	T: 0.857 (180)	T: 0.662 (131)	T: 0.681 (666)	T: 0.803 (808)	T: 0.911 (1204)
	CDSN	G: 0.215 (3517)	G: 0.223 (2682)	G: 0.178 (37)	G: 0.215 (40)	G: 0.160 (33)	G: 0.143 (30)	G: 0.338 (67)	G: 0.319 (312)	G: 0.197 (198)	G: 0.089 (118)
Chr. 12	rs74416240	G: 0.857 (11,182)	G: 0.852 (10,257)	G: 0.899 (187)	G: 0.978 (182)	G: 0.927 (191)	G: 0.900 (189)	G: 0.889 (176)	G: 1.000 (978)	G: 1.000 (1006)	G: 1.000 (1322)
	TCHP	A: 0.143 (1870)	A: 0.148 (1787)	A: 0.101 (21)	A: 0.022 (4)	A: 0.073 (15)	A: 0.100 (21)	A: 0.111 (22)			
	rs925368	A: 0.891 (14,583)	A: 0.859 (10,342)	A: 0.918 (191)	A: 0.978 (182)	A: 0.951 (196)	A: 0.900 (189)	A: 0.894 (177)	A: 1.000 (978)	A: 1.000 (1006)	A: 1.000 (1322)
	GIT2	G: 0.109 (1775)	G: 0.141 (1702)	G: 0.082 (17)	G: 0.022 (4)	G: 0.049 (10)	G: 0.100 (21)	G: 0.106 (21)			
	rs7969300	T: 0.67 (10,959)	T: 0.620 (7464)	T: 0.543 (113)	T: 0.538 (100)	T: 0.471 (97)	T: 0.538 (113)	T: 0.525 (104)	T: 0.093 (91)	T: 0.002 (2)	T: 0.151 (199)
	ATXN2	C: 0.33 (5399)	C: 0.380 (4580)	C: 0.457 (95)	C: 0.462 (86)	C: 0.529 (109)	C: 0.462 (97)	C: 0.475 (94)	C: 0.907 (887)	C: 0.998 (1004)	C: 0.849 (1123)
	rs12231744	C: 0.698 (11,410)	C: 0.650 (7832)	C: 0.591 (123)	C: 0.575 (107)	C: 0.558 (115)	C: 0.614 (129)	C: 0.530 (105)	C: 0.084 (82)	C: 0.003 (3)	C: 0.054 (72)
	NAA25	T: 0.302 (4948)	T: 0.350 (4212)	T: 0.409 (85)	T: 0.425 (79)	T: 0.442 (91)	T: 0.386 (81)	T: 0.470 (93)	T: 0.916 (896)	T: 0.997 (1003)	T: 0.946 (1250)

Values indicate the allele frequency, with the observed numbers in parentheses. Chr, chromosome. JP-Inabe is Japanese in the discovery cohort; JPT is Japanese in Tokyo, Japan; CDX is Chinese Dai in Xishuangbanna, China; CHB is Han Chinese in Beijing, China; CHS is Southern Han Chinese; KHV is Kinh in Ho Chi Minh City, Vietnam.

[a] Allele frequency obtained from the 1000 Genomes Project through the Ensembl genome browser.

functional analyses are required to elucidate the role of *GIT2* in the molecular mechanisms underlying hypo-HDL-cholesterolemia. Allele frequency data from the 1000 Genomes Project indicates that the SNPs in *GIT2* and *TCHP* may be susceptible loci for dyslipidemia specific to East Asian populations.

In the discovery cohort, rs12231744 of *NAA25* exhibited moderate LD with rs7969300 of *ATXN2* ($r^2 = 0.68$). The rs12231744 (T → C, K876R) in the *NAA25* and rs7969300 (C → T, S248N) in *ATXN2* were significantly associated with serum HDL-cholesterol concentration. The protein encoded by *NAA25* is a component of the N-terminal acetyltransferase B complex and is involved in cell cycle progression [20], while ATXN2 protein is involved mRNA translation regulation [21]. These proteins are expressed in various organs and tissues (The Human Protein Atlas). The association of other SNPs in *NAA25* and *ATXN2* with the serum concentrations of LDL-cholesterol or total cholesterol has been determined in a previous GWAS of individuals with European ancestry [13]. Although the functional relevance of the candidate SNPs to serum HDL-cholesterol concentration remains unclear, the association of *NAA25* or *ATXN2* might be attributable to the effect on fluctuations in serum HDL-cholesterol levels in Japanese populations.

The rs34902660 of *SLC17A3* was significantly associated with serum LDL-cholesterol concentration in the present study. The protein encoded by *SLC17A3* is an efflux transporter of intracellular urate and organic anions from the blood into the renal tubule cells. *SLC17A3* is expressed in the liver and kidney (The Human Protein Atlas). Chronic kidney disease is related to the abnormalities of lipid metabolism [22, 23]. Given that *SLC17A3* is related to the alterations in serum lipid concentration via renal regulation, the association of *SLC17A3* with the serum LDL-cholesterol concentration might be attributable to its effect on fluctuations in LDL-cholesterol concentration. The nonsynonymous nucleotide substitution in rs34902660 of *SLC17A3* alters an amino acid at position 161 (C → A, G161V), according to PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml) [24] prediction (PolyPhen score = 0.99), may have deleterious effects on the protein function. This nonsynonymous substitution also results in an amino acid change at position 239 of an *SLC17A3* splice variant (G239 V). According to the SIFT (http://sift.jcvi.org/) [25] prediction, the amino acid substitution may cause a dysfunctional transcript from the splice variant (SIFT score = 0.01). Therefore, rs34902660 of *SLC17A3* could influence serum LDL-cholesterol concentration. According to allele frequency data from the 1000 Genomes Project, the minor allele of this SNP is observed in Japanese and African populations.

The minor allele of rs1042127 in *CDSN* at 6p21.33 may be a risk factor for the incidence of hyper-LDL-cholesterolemia. *CDSN* is a candidate susceptible locus for disorders, such as psoriasis, hypotrichosis, and peeling skin syndrome [26–29]. A nonsynonymous nucleotide substitution T → G at position 408 (S408A) in *CDSN* may cause a dysfunctional protein, as predicted by the PolyPhen-2 (score = 0.928). Therefore, the SNP of *CDSN* may be important for the incidence of hyper-LDL-cholesterolemia. CDSN protein is involved in the shedding of superficial corneocytes from facial skin. Severe psoriasis has been associated with different disorders, including cardiovascular diseases and type 2 diabetes mellitus [30]. Moreover, a previous study has revealed that the mean concentrations of triglycerides, LDL-cholesterol, and total cholesterol in patients with psoriasis were significantly higher than those in controls [31]. Although the relevance of the candidate SNP of *CDSN* in the pathogenesis of hyper-LDL-cholesterolemia remains unclear, the candidate SNP in *CDSN* may be a risk factor for the incidence of this disease.

There were certain limitations in the present study. First, this longitudinal exome-wide association study was conducted in a local Japanese population alone. Thus, the replication of longitudinal exome-wide association studies in other Japanese populations or other ethnic groups is warranted to verify the association of the identified SNPs with specific diseases. Second, the functional relevance of the candidate SNPs to the pathogenesis of the examined diseases remains unclear.

Table 4
Mean serum concentrations of HDL- or LDL-cholesterol in subjects with different genotypes of each candidate SNP in the discovery cohort, based on measurement data of subjects in the latest year.

RefSNP ID	Location[a]	Gene	Genotype	HDL-cholesterol (mmol/L)[b]	*P*-value[c]	LDL-cholesterol (mmol/L)[b]	*P*-value[c]
rs34902660	6: 25,850,874	*SLC17A3*	CC	1.66 ± 0.45	0.4131	3.15 ± 0.80	**0.0005**
			CA	1.64 ± 0.46		3.25 ± 0.80	
			AA	1.68 ± 0.50		3.36 ± 0.84	
rs1042127	6: 31,116,393	*CDSN*	TT	1.66 ± 0.44	0.8840	3.14 ± 0.79	**0.0007**
			TG	1.65 ± 0.45		3.21 ± 0.82	
			GG	1.66 ± 0.47		3.26 ± 0.79	
rs74416240	12: 109,904,793	*TCHP*	GG	1.67 ± 0.45	$\mathbf{5.63 \times 10^{-7}}$	3.16 ± 0.80	0.4527
			GA	1.63 ± 0.44		3.19 ± 0.80	
			AA	1.51 ± 0.37		3.21 ± 0.73	
rs925368	12: 109,953,174	*GIT2*	AA	1.53 ± 0.37	$\mathbf{2.46 \times 10^{-6}}$	3.21 ± 0.73	0.4513
			AG	1.62 ± 0.44		3.19 ± 0.80	
			GG	1.67 ± 0.45		3.16 ± 0.80	
rs7969300	12: 111,555,908	*ATXN2*	TT	1.63 ± 0.43	$\mathbf{4.96 \times 10^{-5}}$	3.17 ± 0.77	0.9005
			TC	1.66 ± 0.45		3.17 ± 0.81	
			CC	1.71 ± 0.48		3.18 ± 0.86	
rs12231744	12: 112,039,251	*NAA25*	CC	1.64 ± 0.44	$\mathbf{1.83 \times 10^{-5}}$	3.16 ± 0.78	0.3972
			CT	1.66 ± 0.45		3.19 ± 0.81	
			TT	1.73 ± 0.48		3.15 ± 0.85	

HDL, high density lipoprotein. LDL, low density lipoprotein.
[a] Location in NCBI build GRCh38.
[b] Quantitative data are means and standard deviations.
[c] Based on Bonferroni's correction, a *P*-value of < 0.008 (0.05/6) was considered statistically significant (by one-way analysis of variance) and is shown in bold.

Therefore, further functional analyses are required to validate the results of this study. Third, the exome-wide association study is a focus genotyping method for whole exons and differs from whole genome sequence analyses which provide a complete coverage of coding and non-coding variants across the entire human genome. The SNPs examined in the present study are only exonic variants.

4. Conclusion

In our longitudinal exome-wide association studies, 82 SNPs were significantly associated with at least one of the six dyslipidemia-related traits (hypertriglyceridemia; hypo-HDL-cholesterolemia; hyper-LDL-cholesterolemia; and serum concentrations of triglycerides, HDL-cholesterol, or LDL-cholesterol). Among these, four and two SNPs were identified as novel susceptibility loci for hypo-HDL-cholesterolemia and hyper-LDL-cholesterolemia, respectively.

5. Materials and methods

5.1. Ethics statement

The study protocol complied with the Declaration of Helsinki and was approved by the Committees on the Ethics of Human Research of Mie University Graduate School of Medicine (Tsu, Mie, Japan) and Inabe General Hospital (Inabe, Mie, Japan). Written informed consent was obtained from all subjects prior to study enrolment in the study.

Table 5
Mean serum concentrations of HDL- or LDL-cholesterol in subjects with different genotypes of each candidate SNP in the replication and entire cohorts.

RefSNP ID	Location[a]	Gene	Genotype	HDL-cholesterol[b]		*P*-value[c]		LDL-cholesterol [b]		*P*-value[c]	
				Replication	Entire[d]	Replication	Entire[d]	Replication	Entire[d]	Replication	Entire[d]
rs34902660	6: 25,850,874	*SLC17A3*	CC	1.36 ± 0.44	1.48 ± 0.47	**0.0003**	0.4746	3.06 ± 0.89	3.10 ± 0.85	0.1349	$\mathbf{5.90 \times 10^{-5}}$
			CA	1.31 ± 0.43	1.47 ± 0.47			3.12 ± 1.00	3.19 ± 0.91		
			AA	1.24 ± 0.50	1.47 ± 0.54			3.23 ± 0.94	3.30 ± 0.89		
rs1042127	6: 31,116,393	*CDSN*	TT	1.36 ± 0.44	1.48 ± 0.47	**0.0059**	0.1543	3.06 ± 0.89	3.09 ± 0.85	0.3712	**0.0008**
			TG	1.33 ± 0.43	1.47 ± 0.47			3.09 ± 0.94	3.15 ± 0.88		
			GG	1.38 ± 0.49	1.51 ± 0.50			3.10 ± 0.93	3.18 ± 0.87		
rs74416240	12: 109,904,793	*TCHP*	GG	1.37 ± 0.45	1.49 ± 0.47	$\mathbf{1.91 \times 10^{-6}}$	$\mathbf{2.93 \times 10^{-6}}$	3.07 ± 0.91	3.11 ± 0.86	0.6093	0.1835
			GA	1.32 ± 0.42	1.46 ± 0.45			3.09 ± 0.91	3.14 ± 0.86		
			AA	1.25 ± 0.38	1.38 ± 0.39			3.09 ± 0.92	3.15 ± 0.82		
rs925368	12: 109,953,174	*GIT2*	AA	1.37 ± 0.45	1.49 ± 0.47	$\mathbf{9.85 \times 10^{-6}}$	$\mathbf{4.35 \times 10^{-5}}$	3.07 ± 0.91	3.11 ± 0.87	0.4452	0.1502
			AG	1.32 ± 0.42	1.45 ± 0.45			3.10 ± 0.92	3.14 ± 0.86		
			GG	1.26 ± 0.38	1.40 ± 0.40			3.04 ± 0.84	3.13 ± 0.78		
rs7969300	12: 111,555,908	*ATXN2*	TT	1.33 ± 0.42	1.45 ± 0.45	$\mathbf{4.13 \times 10^{-7}}$	$\mathbf{5.25 \times 10^{-9}}$	3.12 ± 0.91	3.14 ± 0.85	**0.0044**	0.0249
			TC	1.36 ± 0.44	1.48 ± 0.47			3.05 ± 0.91	3.10 ± 0.87		
			CC	1.40 ± 0.47	1.53 ± 0.50			3.03 ± 0.90	3.10 ± 0.88		
rs12231744	12: 112,039,251	*NAA25*	CC	1.33 ± 0.43	1.45 ± 0.46	$\mathbf{1.88 \times 10^{-6}}$	$\mathbf{8.35 \times 10^{-10}}$	3.11 ± 0.92	3.14 ± 0.85	**0.0038**	0.0584
			CT	1.37 ± 0.44	1.49 ± 0.47			3.04 ± 0.92	3.11 ± 0.88		
			TT	1.40 ± 0.46	1.54 ± 0.49			3.03 ± 0.84	3.08 ± 0.85		

HDL, high density lipoprotein. LDL, low density lipoprotein.
[a] Location in NCBI build GRCh38.
[b] Quantitative data are means and standard deviations (mmol/L).
[c] Based on Bonferroni's correction, a *P*-value of < 0.008 (0.05/6) was considered statistically significant (by one-way analysis of variance) and is shown in bold.
[d] Entire cohort: combined cohort of discovery and replication cohorts.

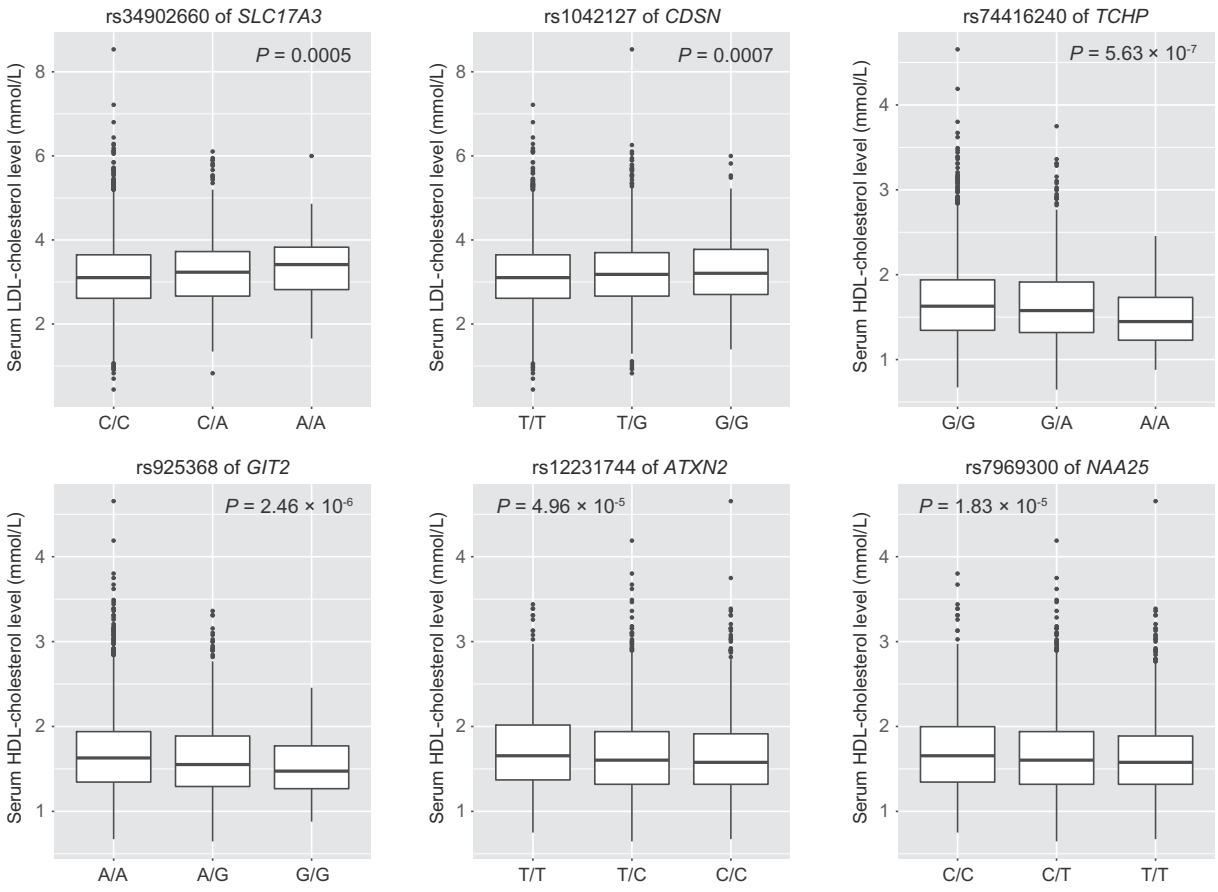

Fig. 5. Box-plots of serum LDL- or HDL-cholesterol levels of subjects with different genotypes of each candidate SNP in the discovery cohort. HDL, high density lipoprotein. LDL, low density lipoprotein.

5.2. Study subjects

In total, 6022 community-dwelling individuals were recruited from among those who visited the Health Care Center at Inabe General Hospital for an annual health checkup and were followed-up (discovery cohort). All subjects had undergone 1–11 medical examinations from April 2003 to March 2014, and the average follow-up period was 5 years. Detailed methods regarding subject recruitment and the collection and storage of medical examination data and genomic DNA samples have been previously described [32]. Cross-sectional data for dyslipidemia-related traits in 8310 Japanese subjects (Gifu Prefectural Tajimi Hospital, Tajimi, Japan; Gifu Prefectural General Medical Center, Gifu, Japan; Japanese Red Cross Nagoya First Hospital, Nagoya, Japan; and Hirosaki University Hospital and Hirosaki Stroke Center, Hirosaki, Japan) were used for the replication studies of candidate SNPs identified from our longitudinal exome-wide association studies. We refer to this cohort as "replication cohort."

Venous blood was collected in the early morning following overnight fasting. The blood samples were centrifuged at 1600 $\times g$ for 15 min at 4 °C, and serum was separated for subsequent analysis. Hypertriglyceridemia was defined as either a serum triglyceride concentration of $\geq$1.69 mmol/L or the current use of antidyslipidemic medications for hypertriglyceridemia. Hypo-HDL-cholesterolemia was defined as a serum HDL-cholesterol concentration of $<$ 1.03 mmol/L. Hyper-LDL-cholesterolemia was defined as a serum LDL-cholesterol concentration of $\geq$3.62 mmol/L or the current treatment with antidyslipidemic agents for hyper-LDL-cholesterolemia. Subjects with dyslipidemia presented with at least one of hypertriglyceridemia, hypo-HDL-cholesterolemia, or hyper-LDL-cholesterolemia, or were receiving antidyslipidemic medications.

In the longitudinal exome-wide association studies for hypertriglyceridemia or hypo-HDL-cholesterolemia, 6022 subjects were examined (3966 controls and 2056 subjects with hypertriglyceridemia or 5324 controls and 698 subjects with hypo-HDL-cholesterolemia, totaling 28,521 examinations per study). In the longitudinal exome-wide association studies for hyper-LDL-cholesterolemia (3251 controls and 2769 subjects with hyper-LDL-cholesterolemia, totaling 28,511 examinations), 6020 subjects were examined. Of these, 540 subjects with both hypertriglyceridemia and hypo-HDL-cholesterolemia and 3808 controls overlapped between the longitudinal exome-wide association studies for these phenotypes. Further, 1334 subjects with both hypertriglyceridemia and hyper-LDL-cholesterolemia and 2530 controls as well as 395 subjects with both hypo-HDL-cholesterolemia and hyper-LDL-cholesterolemia and 2950 controls overlapped between the corresponding studies. The numbers of serum samples examined were 28,038 (5988 subjects) for triglycerides, 28,003 (5988 subjects) for HDL-cholesterol, and 26,831 (5987 subjects) for LDL-cholesterol. The distribution of serum concentrations of triglycerides, HDL-cholesterol, and LDL-cholesterol are presented in Supplementary Fig. S1.

The replication cohort comprised 2685 subjects with hypertriglyceridemia and 4703 controls in the cross-sectional association study for hypertriglyceridemia; 1947 subjects with hypo-HDL-cholesterolemia and 6146 controls in the study for hypo-HDL-cholesterolemia; and 1719 subjects with hyper-LDL-cholesterolemia and 5833 controls in the study for hyper-LDL-cholesterolemia. Of these, 759 subjects with both hypertriglyceridemia and hypo-HDL-cholesterolemia and 4033 controls; 667 subjects with both hypertriglyceridemia and hyper-LDL-cholesterolemia and 3793 controls; and 316 subjects with both hypo-HDL-cholesterolemia and hyper-LDL-cholesterolemia and 4823 controls, overlapped between the corresponding studies.

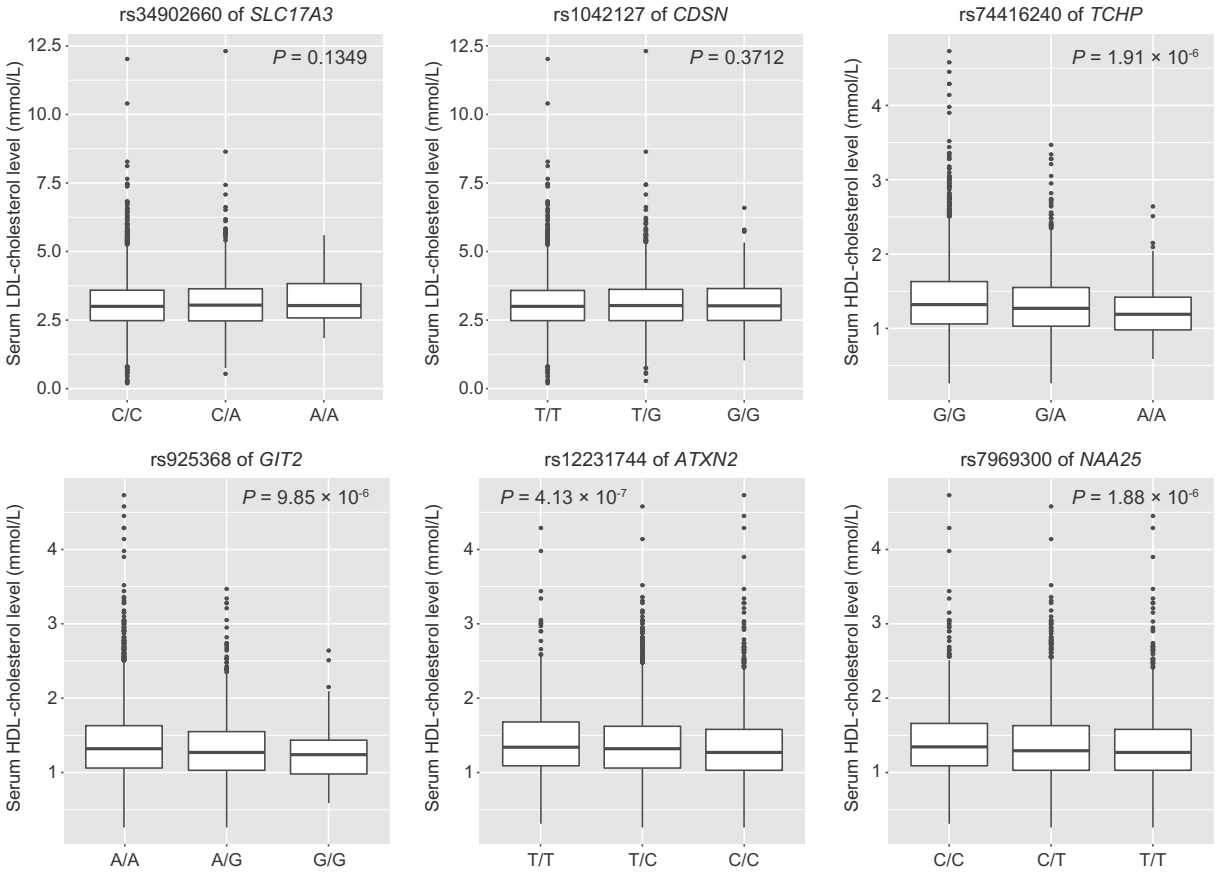

Fig. 6. Box-plots of serum LDL- or HDL-cholesterol levels of subjects with different genotypes of each candidate SNP in the replication cohort. HDL, high density lipoprotein. LDL, low density lipoprotein.

5.3. Longitudinal exome-wide association study

A longitudinal exome-wide association study of the discovery cohort was performed using ~244,000 genetic variants and the longitudinal data from medical examinations. All samples in the discovery cohort were genotyped using Infinium HumanExome-12 ver 1.2 BeadChip and Infinium Exome-24 ver 1.0 BeadChip (Illumina, San Diego, CA, USA). These arrays include putative functional exonic variants selected from > 12,000 individual exome and whole-genome sequences across diverse ethnic populations, including European, African, Chinese, and Hispanic individuals [33].

Quality controls for the genotyping data were performed, and monomorphic sites and the following genetic variants were discarded: variants contained in only one of the exome arrays used (~3.6% of all variants), variants with a call rate of < 97.0%, variants in which the genotype distribution significantly deviated from the Hardy–Weinberg equilibrium ($P < .001$) in controls, variants on the mitochondrial DNA and sex chromosomes, and variants with an MAF of < 0.05. Additionally, the gender specification was examined for each sample, and samples for which the gender designation in the clinical records was inconsistent with the genetic sex were discarded. Cryptic relatedness and duplicate samples were assessed by calculating the identity by descent (IBD); all pairs of DNA samples indicating an IBD of > 0.1875 were inspected, and one sample from each pair was excluded. Principal component analysis (PCA) of SNPs using the EIGENSTRAT method [34] via JMP Genomics version 6.0 (SAS Institute, Cary, NC, USA) was conducted to detect population stratification on a genome-wide scale. No population outliers were identified in PCA. Consequently, 24,691 SNPs among 6022 Japanese individuals (3788 patients with dyslipidemia and 2234 controls) passed the quality control for the longitudinal studies.

Using JMP Genomics, the genotyping data of 6022 individuals were converted into numeric data with three inheritance models (dominant, additive, and recessive). The dominant and recessive models were defined as “AA (0) vs. AB + BB (1)” and “AA + AB (0) vs. BB (1)” (A, major allele; B, minor allele), respectively, whereas the additive model was defined as “AA (0), AB (1), BB (2).” The rearrangement of the longitudinal data was conducted using the R software version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria) [35] via RStudio 1.1.183 [36] and Perl script.

Quantile–quantile plots for the *P*-values in the exome-wide association studies for the clinical parameters related to dyslipidemia in the three inheritance models were performed (Supplementary Figs. S2–S4). The genomic inflation factor (λ) of the *P*-values in the dominant, additive, and recessive models were 1.05 for hypertriglyceridemia, 1.15–1.16 for serum triglyceride concentration, 1.00–1.09 for hypo-HDL-cholesterolemia, 1.09 for serum HDL-cholesterol concentration, 0.98–1.05 for hyper-LDL-cholesterolemia, and 1.00–1.14 for serum LDL-cholesterol concentration.

5.4. Statistical analyses

The association of SNPs with the six dyslipidemia phenotypes was tested using the GEE model [37, 38] with adjustments for age, sex, BMI, and smoking status using the R package “geepack” [39]. Since the prevalence of hypertriglyceridemia, hypo-HDL-cholesterolemia, and hyper-LDL-cholesterolemia was a repeated categorical data (case or control), a binomial distribution was applied for assessing the correlation between the repeated categorical outcomes and SNPs in the GEE method. For instance, the fit with the binomial variance function for the

prevalence of hypertriglyceridemia under a first-order autoregressive (ar1) model for working correlation matrix (correlation of repeated measurements within an individual) using GEE approach in the R package "geepack" is obtained by

geeglm (Prevalence of hypertriglyceridemia~SNP + age + sex + BMI
+smoking status, family = binomial, corstr="ar1")

A Gaussian (normal) distribution was selected for the family with an identity link in the GEE for serum concentrations of triglycerides, HDL-cholesterol, and LDL-cholesterol because the repeated measurements were continuous data. The fit with the serum concentration of triglyceride is obtained by

geeglm (Serum concentration of triglyceride~SNP + age + sex + BMI
+smoking status, family = gaussian, corstr="ar1")

Missing data of response variables or covariates were removed from the GEE analysis. The waves argument was used to specify the ordering of repeated measurements within individuals.

After applying Bonferroni's correction to compensate for multiple comparisons of genotypes with the clinical parameters, the statistical significance of the association was $P < 2.03 \times 10^{-6}$ (0.05/24,691 SNPs) for the three inheritance models. Sitlani et al. [40] have reported that a small effective sample size could increase the probability of generating false positives (type I errors). They have recommended the use of approxdf, a scale of small effective sample size, as follows: approxdf = 2 × MAF × Nindep, where Nindep is the sum of the estimated number of independent observations per subject. The study indicated that an approxdf ≥10 could reduce type I errors. Thus, we estimated the approxdf using the R package "bosswithdf" [40, 41]. To avoid false positive association in small sample sizes, a strict approxdf threshold was applied and SNPs with an approxdf ≤30 were discarded.

5.5. LD estimates

After the SNP data were converted into suitable formats by Perl and R scripts, LDs between pairs of SNPs were estimated using Haploview version 4.2 [42] program. The association of candidate SNPs with dyslipidemia-related phenotypes reported by previous studies was investigated using GRASP (https://grasp.nhlbi.nih.gov/Overview.aspx) [43], GWAS Catalogue (https://www.ebi.ac.uk/) [44], and DisGeNET (http://www.disgenet.org/web/DisGeNET/) [45] databases.

Information regarding the allele frequencies of target SNPs within four ethnic populations (East Asian, South Asian, European, and African) was obtained from the 1000 Genomes Project (http://www.internationalgenome.org/) [46] through the Ensembl genome browser (http://www.ensembl.org) [47]. The categories of the ethnic populations are listed at the following URL: http://www.internationalgenome.org/data-portal/population. To survey the LDs between the candidate SNPs detected in the present study and previously identified dyslipidemia-associated SNPs in JPT from the 1000 Genomes Project, analyses using the LDproxy of LDlink (https://analysistools.nci.nih.gov/LDlink/) [48], which is a web-based application designed to interrogate LD in population groups, and GRASP database were conducted as follows. First, LDs between the candidate SNP and neighboring SNPs in a specific chromosomal region were comprehensively examined using LDproxy. Second, the association between the dyslipidemia-related phenotypes and SNPs that were in LD with the candidate SNP ($r^2 > 0.5$) was investigated using GRASP database. Third, if the SNP in LD with the candidate SNP did not exhibit a significant association ($P > 2.03 \times 10^{-6}$) with the dyslipidemia-related phenotypes, the candidate SNP was identified as a novel susceptibility locus for dyslipidemia.

Conflict of interest statement

The authors declare no competing financial interests.

Authors' contributions

Y. Yasukochi contributed to analysis and interpretation of the data, and to drafting of the manuscript. J. Sakuma and I. Takeuchi contributed to analysis and interpretation of the data as well as revision of the manuscript. K. Kato, M. Oguri, T. Fujimaki, and H. Horibe each contributed to acquisition of the data and revision of the manuscript. Y. Yamada contributed to conception and design of the study, and to acquisition, analysis, interpretation of the data, and revision of the manuscript.

Acknowledgments

The authors deeply thank all the subjects for their participation in the study. This work was supported by Research Grant from Okasan Kato Culture Promotion Foundation (to Y. Yasukochi), the Kurata Grant (Grant number, 1323) awarded by the Hitachi Global Foundation (to Y. Yasukochi, Y. Yamada), CREST (JPMRJCR1302) of the Japan Science and Technology Agency (to Y. Yamada, J. Sakuma, I. Takeuchi), and by Japan Society for the Promotion of ScienceKAKENHI GrantsJP17H00758 (to I. Takeuchi, Y. Yasukochi) and JP15H04772 (to Y. Yamada).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2018.05.015.

References

[1] P.P. Tóth, D. Potter, E.E. Ming, Prevalence of lipid abnormalities in the United States: the National Health and nutrition examination survey 2003–2006, J. Clin. Lipidol. 6 (2012) 325–330, http://dx.doi.org/10.1016/j.jacl.2012.05.002.
[2] Ministry of Health, Labour and Welfare, The National Health and Nutrition Survey in Japan, 2016, http://www.mhlw.go.jp/bunya/kenkou/kenkou_eiyou_chousa.html, (2017) , Accessed date: 1 March 2018.
[3] L. Williams, 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) final report, Circulation 106 (2002) 3143, http://dx.doi.org/10.1001/jama.285.19.2486.
[4] K. Maruyama, K. Hirobe, H. Noda, H. Iso, S. Dohi, T. Terai, S. Fujioka, K. Goto, S. Horie, S. Nakano, Associations between blood lipid profiles and risk of myocardial infarction among Japanese male workers: 3M study, J. Atheroscler. Thromb. 16 (2009) 714–721, http://dx.doi.org/10.5551/jat.547.
[5] J. Sun, Y. Qian, Y. Jiang, J. Chen, J. Dai, G. Jin, J. Wang, Z. Hu, S. Liu, C. Shen, H. Shen, Association of KCTD10, MVK, and MMAB polymorphisms with dyslipidemia and coronary heart disease in Han Chinese population, Lipids Health Dis. 15 (2016) 171, http://dx.doi.org/10.1186/s12944-016-0348-7.
[6] T.M. Teslovich, K. Musunuru, A.V. Smith, A.C. Edmondson, I.M. Stylianou, M. Koseki, J.P. Pirruccello, S. Ripatti, D.I. Chasman, C.J. Willer, C.T. Johansen, S.W. Fouchier, A. Isaacs, G.M. Peloso, M. Barbalic, S.L. Ricketts, J.C. Bis, Y.S. Aulchenko, G. Thorleifsson, M.F. Feitosa, J. Chambers, M. Orho-Melander, O. Melander, T. Johnson, X. Li, X. Guo, M. Li, Y. Shin Cho, M. Jin Go, Y. Jin Kim, J.-Y. Lee, T. Park, K. Kim, X. Sim, R. Twee-Hee Ong, D.C. Croteau-Chonka, L.A. Lange, J.D. Smith, K. Song, J. Hua Zhao, X. Yuan, J. Luan, C. Lamina, A. Ziegler, W. Zhang, R.Y.L. Zee, A.F. Wright, J.C.M. Witteman, J.F. Wilson, G. Willemsen, H.-E. Wichmann, J.B. Whitfield, D.M. Waterworth, N.J. Wareham, G. Waeber, P. Vollenweider, B.F. Voight, V. Vitart, A.G. Uitterlinden, M. Uda, J. Tuomilehto, J.R. Thompson, T. Tanaka, I. Surakka, H.M. Stringham, T.D. Spector, N. Soranzo, J.H. Smit, J. Sinisalo, K. Silander, E.J.G. Sijbrands, A. Scuteri, J. Scott, D. Schlessinger, S. Sanna, V. Salomaa, J. Saharinen, C. Sabatti, A. Ruokonen, I. Rudan, L.M. Rose, R. Roberts, M. Rieder, B.M. Psaty, P.P. Pramstaller, I. Pichler, M. Perola, B.W.J.H. Penninx, N.L. Pedersen, C. Pattaro, A.N. Parker, G. Pare, B.A. Oostra, C.J. O'Donnell, M.S. Nieminen, D.A. Nickerson, G.W. Montgomery, T. Meitinger, R. McPherson, M.I. McCarthy, W. McArdle, D. Masson, N.G. Martin, F. Marroni, M. Mangino, P.K.E. Magnusson, G. Lucas, R. Luben, R.J.F. Loos, M.-L. Lokki, G. Lettre, C. Langenberg, L.J. Launer, E.G. Lakatta, R. Laaksonen, K.O. Kyvik, F. Kronenberg, I.R. König, K.-T. Khaw, J. Kaprio, L.M. Kaplan, A. Johansson, M.-R. Jarvelin, A.C.J.W. Janssens, E. Ingelsson, W. Igl, G. Kees Hovingh, J.-J. Hottenga, A. Hofman, A.A. Hicks, C. Hengstenberg, I.M. Heid, C. Hayward, A.S. Havulinna, N.D. Hastie, T.B. Harris, T. Haritunians, A.S. Hall,

U. Gyllensten, C. Guiducci, L.C. Groop, E. Gonzalez, C. Gieger, N.B. Freimer, L. Ferrucci, J. Erdmann, P. Elliott, K.G. Ejebe, A. Döring, A.F. Dominiczak, S. Demissie, P. Deloukas, E.J.C. de Geus, U. de Faire, G. Crawford, F.S. Collins, Y.I. Chen, M.J. Caulfield, H. Campbell, N.P. Burtt, L.L. Bonnycastle, D.I. Boomsma, S.M. Boekholdt, R.N. Bergman, I. Barroso, S. Bandinelli, C.M. Ballantyne, T.L. Assimes, T. Quertermous, D. Altshuler, M. Seielstad, T.Y. Wong, E.-S. Tai, A.B. Feranil, C.W. Kuzawa, L.S. Adair, H.A. Taylor, I.B. Borecki, S.B. Gabriel, J.G. Wilson, H. Holm, U. Thorsteinsdottir, V. Gudnason, R.M. Krauss, K.L. Mohlke, J.M. Ordovas, P.B. Munroe, J.S. Kooner, A.R. Tall, R.A. Hegele, J.J.P. Kastelein, E.E. Schadt, J.I. Rotter, E. Boerwinkle, D.P. Strachan, V. Mooser, K. Stefansson, M.P. Reilly, N.J. Samani, H. Schunkert, L.A. Cupples, M.S. Sandhu, P.M. Ridker, D.J. Rader, C.M. van Duijn, L. Peltonen, G.R. Abecasis, M. Boehnke, S. Kathiresan, Biological, clinical and population relevance of 95 loci for blood lipids, Nature 466 (2010) 707–713, http://dx.doi.org/10.1038/nature09270.

[7] F.W. Asselbergs, Y. Guo, E.P.A. Van Iperen, S. Sivapalaratnam, V. Tragante, M.B. Lanktree, L.A. Lange, B. Almoguera, Y.E. Appelman, J. Barnard, J. Baumert, A.L. Beitelshees, T.R. Bhangale, Y.D.I. Chen, T.R. Gaunt, Y. Gong, J.C. Hopewell, T. Johnson, M.E. Kleber, T.Y. Langaee, M. Li, Y.R. Li, K. Liu, C.W. McDonough, M.F.L. Meijs, R.P.S. Middelberg, K. Musunuru, C.P. Nelson, J.R. O'Connell, S. Padmanabhan, J.S. Pankow, N. Pankratz, S. Rafelt, R. Rajagopalan, S.P.R. Romaine, N.J. Schork, J. Shaffer, H. Shen, E.N. Smith, S.E. Tischfield, P.J. Van Der Most, J.V. Van Vliet-Ostaptchouk, N. Verweij, K.A. Volcik, L. Zhang, K.R. Bailey, K.M. Bailey, F. Bauer, J.M.A. Boer, P.S. Braund, A. Burt, P.R. Burton, S.G. Buxbaum, W. Chen, R.M. Cooper-Dehoff, L.A. Cupples, J.S. Dejong, C. Delles, D. Duggan, M. Fornage, C.E. Furlong, N. Glazer, J.G. Gums, C. Hastie, M.V. Holmes, T. Illig, S.A. Kirkland, M. Kivimaki, R. Klein, B.E. Klein, C. Kooperberg, K. Kottke-Marchant, M. Kumari, A.Z. Lacroix, L. Mallela, G. Murugesan, J. Ordovas, W.H. Ouwehand, W.S. Post, R. Saxena, H. Scharnagl, P.J. Schreiner, T. Shah, D.C. Shields, D. Shimbo, S.R. Srinivasan, R.P. Stolk, D.I. Swerdlow, H.A. Taylor, E.J. Topol, E. Toskala, J.L. Van Pelt, J. Van Setten, S. Yusuf, J.C. Whittaker, A.H. Zwinderman, S.S. Anand, A.J. Balmforth, G.S. Berenson, C.R. Bezzina, B.O. Boehm, E. Boerwinkle, J.P. Casas, M.J. Caulfield, R. Clarke, J.M. Connell, K.J. Cruickshanks, K.W. Davidson, I.N.M. Day, P.I.W. De Bakker, P.A. Doevendans, A.F. Dominiczak, A.S. Hall, C.A. Hartman, C. Hengstenberg, H.L. Hillege, M.H. Hofker, S.E. Humphries, G.P. Jarvik, J.A. Johnson, B.M. Kaess, S. Kathiresan, W. Koenig, D.A. Lawlor, W. März, O. Melander, B.D. Mitchell, G.W. Montgomery, P.B. Munroe, S.S. Murray, S.J. Newhouse, N.C. Onland-Moret, N. Poulter, B. Psaty, S. Redline, S.S. Rich, J.I. Rotter, H. Schunkert, P. Sever, A.R. Shuldiner, R.L. Silverstein, A. Stanton, B. Thorand, M.D. Trip, M.Y. Tsai, P. Van Der Harst, E. Van Der Schoot, Y.T. Van Der Schouw, W.M.M. Verschuren, H. Watkins, A.A.M. Wilde, B.H.R. Wolffenbuttel, J.B. Whitfield, G.K. Hovingh, C.M. Ballantyne, C. Wijmenga, M.P. Reilly, N.G. Martin, J.G. Wilson, D.J. Rader, N.J. Samani, A.P. Reiner, R.A. Hegele, J.J.P. Kastelein, A.D. Hingorani, P.J. Talmud, H. Hakonarson, C.C. Elbers, B.J. Keating, F. Drenos, Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci, Am. J. Hum. Genet. 91 (2012) 823–838, http://dx.doi.org/10.1016/j.ajhg.2012.08.032.

[8] L.A. Lange, Y. Hu, H. Zhang, C. Xue, E.M. Schmidt, Z.Z. Tang, C. Bizon, E.M. Lange, J.D. Smith, E.H. Turner, G. Jun, H.M. Kang, G. Peloso, P. Auer, K.P. Li, J. Flannick, J. Zhang, C. Fuchsberger, K. Gaulton, C. Lindgren, A. Locke, A. Manning, X. Sim, M.A. Rivas, O.L. Holmen, O. Gottesman, Y. Lu, D. Ruderfer, E.A. Stahl, Q. Duan, Y. Li, P. Durda, S. Jiao, A. Isaacs, A. Hofman, J.C. Bis, A. Correa, M.E. Griswold, J. Jakobsdottir, A.V. Smith, P.J. Schreiner, M.F. Feitosa, Q. Zhang, J.E. Huffman, J. Crosby, C.L. Wassel, R. Do, N. Franceschini, L.W. Martin, J.G. Robinson, T.L. Assimes, D.R. Crosslin, E.A. Rosenthal, M. Tsai, M.J. Rieder, D.N. Farlow, A.R. Folsom, T. Lumley, E.R. Fox, C.S. Carlson, U. Peters, R.D. Jackson, C.M. Van Duijn, A.G. Uitterlinden, D. Levy, J.I. Rotter, H.A. Taylor, V. Gudnason, D.S. Siscovick, M. Fornage, I.B. Borecki, C. Hayward, I. Rudan, Y.E. Chen, E.P. Bottinger, R.J.F. Loos, P. Strom, K. Hveem, M. Boehnke, L. Groop, M. McCarthy, T. Meitinger, C.M. Ballantyne, S.B. Gabriel, C.J. O'Donnell, W.S. Post, K.E. North, A.P. Reiner, E. Boerwinkle, B.M. Psaty, D. Altshuler, S. Kathiresan, D.Y. Lin, G.P. Jarvik, L.A. Cupples, C. Kooperberg, J.G. Wilson, D.A. Nickerson, G.R. Abecasis, S.S. Rich, R.P. Tracy, C.J. Willer, Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol, Am. J. Hum. Genet. 94 (2014) 233–245, http://dx.doi.org/10.1016/j.ajhg.2014.01.010.

[9] J.A. Hubacek, V. Adamkova, V. Lanska, D. Dlouha, Polygenic hypercholesterolemia: examples of GWAS results and their replication in the Czech-Slavonic population, Physiol. Res. 66 (2017) S101–S111.

[10] X. Lu, J. Li, H. Li, Y. Chen, L. Wang, M. He, Y. Wang, L. Sun, Y. Hu, J. Huang, F. Wang, X. Liu, S. Chen, K. Yu, X. Yang, Z. Mo, X. Lin, T. Wu, D. Gu, Coding - sequence variants are associated with blood lipid levels in 14,473 Chinese, Hum. Mol. Genet. 25 (2016) 4107–4116, http://dx.doi.org/10.1093/hmg/ddw261.

[11] C.S. Carlson, M.A. Eberle, M.J. Rieder, Q. Yi, L. Kruglyak, D.A. Nickerson, Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium, Am. J. Hum. Genet. 74 (2004) 106–120, http://dx.doi.org/10.1086/381000.

[12] C.A. Anderson, G. Boucher, C.W. Lees, A. Franke, M. D'Amato, K.D. Taylor, J.C. Lee, P. Goyette, M. Imielinski, A. Latiano, C. Lagacé, R. Scott, L. Amininejad, S. Bumpstead, L. Baidoo, R.N. Baldassano, M. Barclay, T.M. Bayless, S. Brand, C. Büning, J.F. Colombel, L.A. Denson, M. De Vos, M. Dubinsky, C. Edwards, D. Ellinghaus, R.S.N. Fehrmann, J.A.B. Floyd, T. Florin, D. Franchimont, L. Franke, M. Georges, J. Glas, N.L. Glazer, S.L. Guthery, T. Haritunians, N.K. Hayward, J.P. Hugot, G. Jobin, D. Laukens, I. Lawrance, M. Lémann, A. Levine, C. Libioulle, E. Louis, D.P. McGovern, M. Milla, G.W. Montgomery, K.I. Morley, C. Mowat, A. Ng, W. Newman, R.A. Ophoff, L. Papi, O. Palmieri, L. Peyrin-Biroulet, J. Panés, A. Phillips, N.J. Prescott, D.D. Proctor, R. Roberts, R. Russell, P. Rutgeerts, J. Sanderson, M. Sans, P. Schumm, F. Seibold, Y. Sharma, L.A. Simms, M. Seielstad, A.H. Steinhart, S.R. Targan, L.H. Van Den Berg, M. Vatn, H. Verspaget, T. Walters, C. Wijmenga, D.C. Wilson, H.J. Westra, R.J. Xavier, Z.Z. Zhao, C.Y. Ponsioen, V. Andersen, L. Torkvist, M. Gazouli, N.P. Anagnou, T.H. Karlsen, L. Kupcinskas, J. Sventoraityte, J.C. Mansfield, S. Kugathasan, M.S. Silverberg, J. Halfvarson, J.I. Rotter, C.G. Mathew, A.M. Griffiths, R. Gearry, T. Ahmad, S.R. Brant, M. Chamaillard, J. Satsangi, J.H. Cho, S. Schreiber, M.J. Daly, J.C. Barrett, M. Parkes, V. Annese, H. Hakonarson, G. Radford-Smith, R.H. Duerr, S. Vermeire, R.K. Weersma, J.D. Rioux, Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47, Nat. Genet. 43 (2011) 246–252, http://dx.doi.org/10.1038/ng.764.

[13] T.M. Teslovich, K. Musunuru, A.V. Smith, A.C. Edmondson, I.M. Stylianou, M. Koseki, J.P. Pirruccello, S. Ripatti, D.I. Chasman, C.J. Willer, C.T. Johansen, S.W. Fouchier, A. Isaacs, G.M. Peloso, M. Barbalic, S.L. Ricketts, J.C. Bis, Y.S. Aulchenko, G. Thorleifsson, M.F. Feitosa, J. Chambers, M. Orho-Melander, O. Melander, T. Johnson, X. Li, X. Guo, M. Li, Y. Shin Cho, M. Jin Go, Y. Jin Kim, J.Y. Lee, T. Park, K. Kim, X. Sim, R. Twee-Hee Ong, D.C. Croteau-Chonka, L.A. Lange, J.D. Smith, K. Song, J. Hua Zhao, X. Yuan, J. Luan, C. Lamina, A. Ziegler, W. Zhang, R.Y.L. Zee, A.F. Wright, J.C.M. Witteman, J.F. Wilson, G. Willemsen, H.E. Wichmann, J.B. Whitfield, D.M. Waterworth, N.J. Wareham, G. Waeber, P. Vollenweider, B.F. Voight, V. Vitart, A.G. Uitterlinden, M. Uda, J. Tuomilehto, J.R. Thompson, T. Tanaka, I. Surakka, H.M. Stringham, T.D. Spector, N. Soranzo, J.H. Smit, J. Sinisalo, K. Silander, E.J.G. Sijbrands, A. Scuteri, J. Scott, D. Schlessinger, S. Sanna, V. Salomaa, J. Saharinen, C. Sabatti, A. Ruokonen, I. Rudan, L.M. Rose, R. Roberts, M. Rieder, B.M. Psaty, P.P. Pramstaller, I. Pichler, M. Perola, B.W.J.H. Penninx, N.L. Pedersen, C. Pattaro, A.N. Parker, G. Pare, B.A. Oostra, C.J. O'donnell, M.S. Nieminen, D.A. Nickerson, G.W. Montgomery, T. Meitinger, R. Mcpherson, M.I. Mccarthy, W. Mcardle, D. Masson, N.G. Martin, F. Marroni, M. Mangino, P.K.E. Magnusson, G. Lucas, R. Luben, R.J.F. Loos, M.L. Lokki, G. Lettre, C. Langenberg, L.J. Launer, E.G. Lakatta, R. Laaksonen, K.O. Kyvik, F. Kronenberg, I.R. König, K.T. Khaw, J. Kaprio, L.M. Kaplan, Å. Johansson, M.R. Jarvelin, A. Cecile, E. Ingelsson, W. Igl, G. Kees Hovingh, J.J. Hottenga, A. Hofman, A.A. Hicks, C. Hengstenberg, I.M. Heid, C. Hayward, A.S. Havulinna, N.D. Hastie, T.B. Harris, T. Haritunians, A.S. Hall, U. Gyllensten, C. Guiducci, L.C. Groop, E. Gonzalez, C. Gieger, N.B. Freimer, L. Ferrucci, J. Erdmann, P. Elliott, K.G. Ejebe, A. Döring, A.F. Dominiczak, S. Demissie, P. Deloukas, E.J.C. De Geus, U. De Faire, G. Crawford, F.S. Collins, Y.D.I. Chen, M.J. Caulfield, H. Campbell, N.P. Burtt, L.L. Bonnycastle, D.I. Boomsma, S.M. Boekholdt, R.N. Bergman, I. Barroso, S. Bandinelli, C.M. Ballantyne, T.L. Assimes, T. Quertermous, D. Altshuler, M. Seielstad, T.Y. Wong, E.S. Tai, A.B. Feranil, C.W. Kuzawa, L.S. Adair, H.A. Taylor, I.B. Borecki, S.B. Gabriel, J.G. Wilson, H. Holm, U. Thorsteinsdottir, V. Gudnason, R.M. Krauss, K.L. Mohlke, J.M. Ordovas, P.B. Munroe, J.S. Kooner, A.R. Tall, R.A. Hegele, J.J.P. Kastelein, E.E. Schadt, J.I. Rotter, E. Boerwinkle, D.P. Strachan, V. Mooser, K. Stefansson, M.P. Reilly, N.J. Samani, H. Schunkert, L.A. Cupples, M.S. Sandhu, P.M. Ridker, D.J. Rader, C.M. Van Duijn, L. Peltonen, G.R. Abecasis, M. Boehnke, S. Kathiresan, Biological, clinical and population relevance of 95 loci for blood lipids, Nature 466 (2010) 707–713, http://dx.doi.org/10.1038/nature09270.

[14] M.J. Barber, L.M. Mangravite, C.L. Hyde, D.I. Chasman, J.D. Smith, C.A. McCarty, X. Li, R.A. Wilke, M.J. Rieder, P.T. Williams, P.M. Ridker, A. Chatterjee, J.I. Rotter, D.A. Nickerson, M. Stephens, R.M. Krauss, Genome-side association of lipid-lowering response to statins in combined study populations, PLoS One 5 (2010) e9763, , http://dx.doi.org/10.1371/journal.pone.0009763.

[15] S. Kathiresan, C.J. Willer, G.M. Peloso, S. Demissie, K. Musunuru, E.E. Schadt, L. Kaplan, D. Bennett, Y. Li, T. Tanaka, B.F. Voight, L.L. Bonnycastle, A.U. Jackson, G. Crawford, A. Surti, C. Guiducci, N.P. Burtt, S. Parish, R. Clarke, D. Zelenika, K.A. Kubalanza, M.A. Morken, L.J. Scott, H.M. Stringham, P. Galan, A.J. Swift, J. Kuusisto, R.N. Bergman, J. Sundvall, M. Laakso, L. Ferrucci, P. Scheet, S. Sanna, M. Uda, Q. Yang, K.L. Lunetta, J. Dupuis, P.I.W. De Bakker, C.J. O'Donnell, J.C. Chambers, J.S. Kooner, S. Hercberg, P. Meneton, E.G. Lakatta, A. Scuteri, D. Schlessinger, J. Tuomilehto, F.S. Collins, L. Groop, D. Altshuler, R. Collins, G.M. Lathrop, O. Melander, V. Salomaa, L. Peltonen, M. Orho-Melander, J.M. Ordovas, M. Boehnke, G.R. Abecasis, K.L. Mohlke, L.A. Cupples, Common variants at 30 loci contribute to polygenic dyslipidemia, Nat. Genet. 41 (2009) 56–65, http://dx.doi.org/10.1038/ng.291.

[16] C.J. Willer, E.M. Schmidt, S. Sengupta, G.M. Peloso, S. Gustafsson, S. Kanoni, A. Ganna, J. Chen, M.L. Buchkovich, S. Mora, J.S. Beckmann, J.L. Bragg-Gresham, H.Y. Chang, A. Demirkan, H.M. Den Hertog, R. Do, L.A. Donnelly, G.B. Ehret, T. Esko, M.F. Feitosa, T. Ferreira, K. Fischer, P. Fontanillas, R.M. Fraser, D.F. Freitag, D. Gurdasani, K. Heikkilä, E. Hyppönen, A. Isaacs, A.U. Jackson, Å. Johansson, T. Johnson, M. Kaakinen, J. Kettunen, M.E. Kleber, X. Li, J. Luan, L.P. Lyytikäinen, P.K.E. Magnusson, M. Mangino, E. Mihailov, M.E. Montasser, M. Müller-Nurasyid, I.M. Nolte, J.R. O'Connell, C.D. Palmer, M. Perola, A.K. Petersen, S. Sanna, R. Saxena, S.K. Service, S. Shah, D. Shungin, C. Sidore, C. Song, R.J. Strawbridge, I. Surakka, T. Tanaka, T.M. Teslovich, G. Thorleifsson, E.G. Van Den Herik, B.F. Voight, K.A. Volcik, L.L. Waite, A. Wong, Y. Wu, W. Zhang, D. Absher, G. Asiki, I. Barroso, L.F. Been, J.L. Bolton, L.L. Bonnycastle, P. Brambilla, M.S. Burnett, G. Cesana, M. Dimitriou, A.S.F. Doney, A. Döring, P. Elliott, S.E. Epstein, G.I. Eyjolfsson, B. Gigante, M.O. Goodarzi, H. Grallert, M.L. Gravito, C.J. Groves, G. Hallmans, A.L. Hartikainen, C. Hayward, D. Hernandez, A.A. Hicks, H. Holm, Y.J. Hung, T. Illig, M.R. Jones, P. Kaleebu, J.J.P. Kastelein, K.T. Khaw, E. Kim, N. Klopp, P. Komulainen, M. Kumari, C. Langenberg, T. Lehtimäki, S.Y. Lin, J. Lindström, R.J.F. Loos, F. Mach, W.L. McArdle, C. Meisinger, B.D. Mitchell, G. Müller, R. Nagaraja, N. Narisu, T.V.M. Nieminen, R.N. Nsubuga, I. Olafsson, K.K. Ong, A. Palotie, T. Papamarkou, C. Pomilla, A. Pouta, D.J. Rader, M.P. Reilly, P.M. Ridker, F. Rivadeneira, I. Rudan, A. Ruokonen, N. Samani, H. Scharnagl, J. Seeley, K. Silander, A. Stancáková, K. Stirrups, A.J. Swift, L. Tiret,

A.G. Uitterlinden, L.J. Van Pelt, S. Vedantam, N. Wainwright, C. Wijmenga, S.H. Wild, G. Willemsen, T. Wilsgaard, J.F. Wilson, E.H. Young, J.H. Zhao, L.S. Adair, D. Arveiler, T.L. Assimes, S. Bandinelli, F. Bennett, M. Bochud, B.O. Boehm, D.I. Boomsma, I.B. Borecki, S.R. Bornstein, P. Bovet, M. Burnier, H. Campbell, A. Chakravarti, J.C. Chambers, Y.D.I. Chen, F.S. Collins, R.S. Cooper, J. Danesh, G. Dedoussis, U. De Faire, A.B. Feranil, J. Ferrières, L. Ferrucci, N.B. Freimer, C. Gieger, L.C. Groop, V. Gudnason, U. Gyllensten, A. Hamsten, T.B. Harris, A. Hingorani, J.N. Hirschhorn, A. Hofman, G.K. Hovingh, C.A. Hsiung, S.E. Humphries, S.C. Hunt, K. Hveem, C. Iribarren, M.R. Järvelin, A. Jula, M. Kähönen, J. Kaprio, A. Kesäniemi, M. Kivimaki, J.S. Kooner, P.J. Koudstaal, R.M. Krauss, D. Kuh, J. Kuusisto, K.O. Kyvik, M. Laakso, T.A. Lakka, L. Lind, C.M. Lindgren, N.G. Martin, W. März, M.I. McCarthy, C.A. McKenzie, P. Meneton, A. Metspalu, L. Moilanen, A.D. Morris, P.B. Munroe, I. Njølstad, N.L. Pedersen, C. Power, P.P. Pramstaller, J.F. Price, B.M. Psaty, T. Quertermous, R. Rauramaa, D. Saleheen, V. Salomaa, D.K. Sanghera, J. Saramies, P.E.H. Schwarz, W.H.H. Sheu, A.R. Shuldiner, A. Siegbahn, T.D. Spector, K. Stefansson, D.P. Strachan, B.O. Tayo, E. Tremoli, J. Tuomilehto, M. Uusitupa, C.M. Van Duijn, P. Vollenweider, L. Wallentin, N.J. Wareham, J.B. Whitfield, B.H.R. Wolffenbuttel, J.M. Ordovas, E. Boerwinkle, C.N.A. Palmer, U. Thorsteinsdottir, D.I. Chasman, J.I. Rotter, P.W. Franks, S. Ripatti, L.A. Cupples, M.S. Sandhu, S.S. Rich, M. Boehnke, P. Deloukas, S. Kathiresan, K.L. Mohlke, E. Ingelsson, G.R. Abecasis, Discovery and refinement of loci associated with lipid levels, Nat. Genet. 45 (2013) 1274–1285, http://dx.doi.org/10.1038/ng.2797.

[17] D. Zabaneh, D.J. Balding, A genome-wide association study of the metabolic syndrome in Indian Asian men, PLoS One 5 (2010) e11961, , http://dx.doi.org/10.1371/journal.pone.0011961.

[18] R.J. Hoefen, The multifunctional GIT family of proteins, J. Cell Sci. 119 (2006) 1469–1475, http://dx.doi.org/10.1242/jcs.02925.

[19] R.P.H. Huijbregts, L. Topalof, V.A. Bankaitis, Lipid metabolism and regulation of membrane trafficking, Traffic 1 (2000) 195–202, http://dx.doi.org/10.1034/j.1600-0854.2000.010301.x.

[20] K.K. Starheim, T. Arnesen, D. Gromyko, A. Ryningen, J.E. Varhaug, J.R. Lillehaug, Identification of the human N α -acetyltransferase complex B (hNatB): a complex important for cell-cycle progression, Biochem. J. 415 (2008) 325–331, http://dx.doi.org/10.1042/BJ20080658.

[21] H.T. Orr, Cell biology of spinocerebellar ataxia, J. Cell Biol. 197 (2012) 167–177, http://dx.doi.org/10.1083/jcb.201105092.

[22] V. Tsimihodimos, Z. Mitrogianni, M. Elisaf, Dyslipidemia associated with chronic kidney disease, Open Cardiovasc. Med. J. 5 (2011) 41–48, http://dx.doi.org/10.2174/1874192401105010041.

[23] R. Trevisan, A.R. Dodesini, G. Lepore, Lipids and renal disease, J. Am. Soc. Nephrol. 17 (2006) S145–S147, http://dx.doi.org/10.1681/ASN.2005121320.

[24] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249, http://dx.doi.org/10.1038/nmeth0410-248.

[25] P.C. Ng, S. Henikoff, Predicting deleterious amino acid substitutions, Genome Res. 11 (2001) 863–874, http://dx.doi.org/10.1101/gr.176601.

[26] R.P. Nair, P.E. Stuart, I. Nistor, R. Hiremagalore, N.V.C. Chia, S. Jenisch, M. Weichenthal, G.R. Abecasis, H.W. Lim, E. Christophers, J.J. Voorhees, J.T. Elder, Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene, Am. J. Hum. Genet. 78 (2006) 827–851, http://dx.doi.org/10.1086/503821.

[27] A. Mallet, M. Kypriotou, K. George, E. Leclerc, D. Rivero, J. Mazereeuw-Hautier, G. Serre, M. Huber, N. Jonca, D. Hohl, Identification of the first nonsense CDSN mutation with expression of a truncated protein causing peeling skin syndrome type B, Br. J. Dermatol. 169 (2013) 1322–1325, http://dx.doi.org/10.1111/bjd.12593.

[28] S. Israeli, H. Zamir, O. Sarig, R. Bergman, E. Sprecher, Inflammatory peeling skin syndrome caused by a mutation in CDSN encoding corneodesmosin, J. Invest. Dermatol. 131 (2011) 779–781, http://dx.doi.org/10.1038/jid.2010.363.

[29] E. Levy-Nissenbaum, R.C. Betz, M. Frydman, M. Simon, H. Lahat, T. Bakhan, B. Goldman, A. Bygum, M. Pierick, A.M. Hillmer, N. Jonca, J. Toribio, R. Kruse, G. Dewald, S. Cichon, C. Kubisch, M. Guerrin, G. Serre, M.M. Nöthen, E. Pras, Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin, Nat. Genet. 34 (2003) 151–153, http://dx.doi.org/10.1038/ng1163.

[30] K. Abuabara, R.S. Azfar, D.B. Shin, A.L. Neimann, A.B. Troxel, J.M. Gelfand, Cause-specific mortality in patients with severe psoriasis: a population-based cohort study in the U.K, Br. J. Dermatol. 163 (2010) 586–592, http://dx.doi.org/10.1111/j.1365-2133.2010.09941.x.

[31] N. Solak Tekin, I.O. Tekin, F. Barut, E. Yilmaz Sipahi, Accumulation of oxidized low-density lipoprotein in psoriatic skin and changes of plasma lipid levels in psoriatic patients, Mediat. Inflamm. (2007) (2007) 78454, http://dx.doi.org/10.1155/2007/78454.

[32] Y. Yamada, K. Matsui, I. Takeuchi, M. Oguri, T. Fujimaki, Association of genetic variants with hypertension in a longitudinal population-based genetic epidemiological study, Int. J. Mol. Med. 35 (2015) 1189–1198, http://dx.doi.org/10.3892/ijmm.2015.2151.

[33] M.L. Grove, B. Yu, B.J. Cochran, T. Haritunians, J.C. Bis, K.D. Taylor, M. Hansen, I.B. Borecki, L.A. Cupples, M. Fornage, V. Gudnason, T.B. Harris, S. Kathiresan, R. Kraaij, L.J. Launer, D. Levy, Y. Liu, T. Mosley, G.M. Peloso, B.M. Psaty, S.S. Rich, F. Rivadeneira, D.S. Siscovick, A.V. Smith, A. Uitterlinden, C.M. van Duijn, J.G. Wilson, C.J. O'Donnell, J.I. Rotter, E. Boerwinkle, J. Korn, F. Kuruvilla, S. McCarroll, A. Wysoker, J. Nemesh, M. Ritchie, R. Liu, B. Carvalho, R. Irizarry, B. Psaty, C. O'Donnell, V. Gudnason, K. Lunetta, A. Folsom, S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. Ferreira, V. Gudnason, G. Sigurdsson, H. Nissen, S. Humphries, S. Thorlacius, G. Olafsdottir, L. Tryggvadottir, S. Neuhausen, J. Jonasson, J. Goldstein, A. Crenshaw, J. Carey, G. Grant, J. Maguire, T. Harris, L. Launer, G. Eiriksdottir, O. Kjartansson, P. Jonsson, I.T. ARIC, D. Siscovick, T. Raghunathan, I. King, S. Weinmann, K. Wicklund, L. Fried, N. Borhani, P. Enright, C. Furberg, J. Gardin, G. Tell, L. Fried, B. Hermanson, T. Manolio, A. Newman, G. Friedman, G. Cutter, R. Donahue, G. Hughes, S. Hulley, G. Cutter, G. Burke, A. Dyer, G. Friedman, J. Hilner, D. Bild, D. Bluemke, G. Burke, R. Detrano, A.D. Roux, M. Higgins, M. Province, G. Heiss, J. Eckfeldt, R. Ellison, T. Dawber, G. Meadors, F.M. Jr, S. Park, B. Goodpaster, E. Strotmeyer, N. de Rekeneire, T. Harris, H.T. Jr, J. Wilson, D. Jones, D. Sarpong, A. Srinivasan, A. Hofman, D. Grobbee, P. de Jong, F. van den Ouweland, A. Hofman, M. Breteler, C. van Duijn, G. Krestin, H. Pols, A. Hofman, M. Breteler, C. van Duijn, H. Janssen, G. Krestin, A. Hofman, C. van Duijn, O. Franco, M. Ikram, H. Janssen, P. Mills, X. Liu, X. Jian, E. Boerwinkle, Best practices and joint calling of the HumanExome BeadChip: the CHARGE consortium, PLoS One 8 (2013) e68095, , http://dx.doi.org/10.1371/journal.pone.0068095.

[34] A.L. Price, N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies, Nat. Genet. 38 (2006) 904–909, http://dx.doi.org/10.1038/ng1847.

[35] R Core Team, R: A Language and Environment for Statistical Computing, https://www.r-project.org/, (2016).

[36] RStudio Team, RStudio: Integrated Development Environment for R, RStudio, Inc, 2015, http://www.rstudio.com/.

[37] K.Y. Liang, S.L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986) 13–22, http://dx.doi.org/10.1093/biomet/73.1.13.

[38] J.A. Hanley, A. Negassa, M.D. Edwardes, J.E. Forrester, Statistical analysis of correlated data using generalized estimating equations: an orientation, Am. J. Epidemiol. 157 (2003) 364–375, http://dx.doi.org/10.1093/aje/kwf215.

[39] U. Halekoh, S. Højsgaard, J. Yan, The R package geepack for generalized estimating equations, J. Stat. Softw. 15 (2006) 1–11.

[40] C.M. Sitlani, K.M. Rice, T. Lumley, B. McKnight, L.A. Cupples, C.L. Avery, R. Noordam, B.H.C. Stricker, E.A. Whitsel, B.M. Psaty, Generalized estimating equations for genome-wide association studies using longitudinal phenotype data, Stat. Med. 34 (2015) 118–130, http://dx.doi.org/10.1002/sim.6323.

[41] A. Voorman, K. Rice, T. Lumley, Fast computation for genome-wide association studies using boosted one-step statistics, Bioinformatics 28 (2012) 1818–1822, http://dx.doi.org/10.1093/bioinformatics/bts291.

[42] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, Bioinformatics 21 (2005) 263–265, http://dx.doi.org/10.1093/bioinformatics/bth457.

[43] R. Leslie, C.J. O'Donnell, A.D. Johnson, GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database, Bioinformatics 30 (2014) 185–194, http://dx.doi.org/10.1093/bioinformatics/btu273.

[44] J. MacArthur, E. Bowler, M. Cerezo, L. Gil, P. Hall, E. Hastings, H. Junkins, A. McMahon, A. Milano, J. Morales, Z. MayPendlington, D. Welter, T. Burdett, L. Hindorff, P. Flicek, F. Cunningham, H. Parkinson, The new NHGRI-EBI catalog of published genome-wide association studies (GWAS catalog), Nucleic Acids Res. 45 (2017) D896–D901, http://dx.doi.org/10.1093/nar/gkw1133.

[45] J. Piñero, N. Queralt-Rosinach, À. Bravo, J. Deu-Pons, A. Bauer-Mehren, M. Baron, F. Sanz, L.I. Furlong, DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes, Database 2015 (2015) bav028, , http://dx.doi.org/10.1093/database/bav028.

[46] The 1000 Genomes Project Consortium, A map of human genome variation from population-scale sequencing, Nature 467 (2010) 1061–1073, http://dx.doi.org/10.1038/nature09534.

[47] D.R. Zerbino, P. Achuthan, W. Akanni, M.R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C.G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O.G. Izuogu, S.H. Janacek, T. Juettemann, J.K. To, M.R. Laird, I. Lavidas, Z. Liu, J.E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D.N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C.K. Ong, A. Parker, M. Patricio, H.S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S.E. Hunt, M. Kostadima, N. Langridge, F.J. Martin, M. Muffato, E. Perry, M. Ruffier, D.M. Staines, S.J. Trevanion, B.L. Aken, F. Cunningham, A. Yates, P. Flicek, Ensembl 2018, Nucleic Acids Res. 46 (2018) D754–D761 https://doi.org/10.1093/nar/gkx1098.

[48] M.J. Machiela, S.J. Chanock, LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants, Bioinformatics 31 (2015) 3555–3557, http://dx.doi.org/10.1093/bioinformatics/btv402.