Association between \textit{CETP}, \textit{MLXIPL}, and \textit{TOMM40} polymorphisms and serum lipid levels in a Latvian population

I. Radovica, D. Fridmanis, I. Silamikelis, L. Nikitina-Zake, J. Klovins *

\textit{Latvian Biomedical Research and Study Centre, Latvia, Riga}

\section*{Article Info}

\textbf{Article history:} \\
Received 30 April 2014 \\
Revised 24 July 2014 \\
Accepted 25 July 2014 \\
Available online 20 August 2014

\textbf{Keywords:} \\
Association study \\
Dyslipidemia \\
High-density lipoprotein–cholesterol \\
Total cholesterol \\
Low-density lipoprotein–cholesterol \\
Triglycerides

\section*{Abstract}

\textbf{Background:} Abnormal lipid levels are considered one of the most significant risk factors for atherosclerosis and coronary artery disease, two of the main causes of death worldwide. Apart from monogenic cases of hypercholesterolemia, most of the common dyslipidemias are caused by a number of low-impact polymorphisms. It has recently been reported that frequent polymorphisms at a large number of loci are significantly associated with one or more blood lipid parameters in many populations. Identifying these associations in different populations and estimating the possible interactions between genetic models are necessary to explain the underlying genetic architecture of the associated loci and their ultimate impact on lipid-associated traits.

\textbf{Methods}: We estimated the association between 144 common single-nucleotide polymorphisms (SNPs) from published genome-wide association studies and the levels of total cholesterol, low- and high-density lipoprotein–cholesterol, and triglycerides in 1273 individuals from the Genome Database of the Latvian Population. We analyzed a panel of 144 common SNPs with Illumina GoldenGate Genotyping Assays on the Illumina BeadXpress System.

\textbf{Results}: Ten SNPs at the \textit{CETP} locus and two at the \textit{MLXIPL} locus were associated with reduced high-density lipoprotein–cholesterol levels; one SNP at the \textit{TOMM40} locus was associated with increased low-density lipoprotein–cholesterol; and four SNPs at the \textit{MLXIPL} locus were associated with increased log triglyceride levels. There was also a significant correlation between the number of risk alleles and all the...
Introduction

Dyslipidemia is a disorder of lipid and lipoprotein metabolism, including overproductivity, deficiencies, or other changes in lipid metabolism, which results in abnormal amounts of lipids and lipoproteins in the blood. Dyslipidemia can manifest as an elevation of the total cholesterol (TC), low-density lipoprotein (LDL)–cholesterol, or triglyceride (TG) concentration, or a reduction in the high-density lipoprotein (HDL)–cholesterol concentration in the blood (Goldberg, 2013). Abnormal lipid levels are one of the essential causes of atherosclerosis and coronary artery disease (CAD) (Vance and Vance, 2008; World Health Organization, H.G.P., 1999). When lipoproteins in the plasma, and hence in the arterial wall, fall below a certain threshold level, atherogenesis does not occur. However, the absolute value of this threshold varies among individuals and may be very low in subjects with genetic and/or environmental factors that render their arteries highly susceptible to atherogenic processes (Vance and Vance, 2008; World Health Organization, H.G.P., 1999). Polygenic hypercholesterolemia is characterized by moderate elevations in LDL–cholesterol (3.63–7.77 mmol/L), with serum TG concentrations within the reference range. Like more extreme monogenic diseases, polygenic dyslipidemia is also associated with CAD. Although the main causes of dyslipidemia are high fat intake (particularly saturated fats) and obesity, genetic factors are also considered important, because many individuals vary in their responses to dietary cholesterol. However, there is no clear pattern of inheritance, and a combination of several genetic variants is generally required for this type of hypercholesterolemia (Durrington, 2003). Epidemiological studies have established that a reduction in plasma TC reduces the risk of coronary heart disease (CHD) (Bahgavan, 2002), indicating the importance of characterizing the most strongly associated SNPs with the main risk factors for elevated blood lipid levels. This characterization should facilitate the early diagnosis of these risk factors, before the development of CAD symptoms, and the administration of appropriate treatment. Recent genome-wide association studies (GWAS) and other human genetic studies have localized many common SNPs and many loci that influence the levels of different blood lipids, including previously known loci that are potentially involved in lipid metabolism (Aulchenko et al., 2009; Braun et al.; Burkhardt et al., 2008; Chasman et al., 2008; Edmondson et al.; Heid et al., 2008; Hiura et al., 2008; Igl et al.; Johansen et al.; Kathiresan et al., 2008; Kathiresan et al., 2009; Lanktree et al., 2009; Suchindran et al.; Wallace et al., 2008; Waterworth et al.; Willer et al., 2008). More than 40 loci have been associated with the levels of one or more blood lipid parameters, and contain many SNPs, but not all of these associations have been confirmed in other populations (Aulchenko et al., 2009; Burkhardt et al., 2008; Chasman et al., 2008; Heid et al., 2008; Hiura et al., 2008; Igl et al.; Kathiresan et al., 2008; Kathiresan et al., 2009; Suchindran et al.; Wallace et al., 2008; Waterworth et al.; Willer et al., 2008). It has been demonstrated that a number of missense mutations at some of these loci can cause different types of monogenic hypercholesterolemia. For example, mutations in the ABCA1 gene cause familial HDL deficiency, or Tangier disease (Bhagavan, 2002). Similarly, mutations in the APOA5, APOB, and APOE genes cause different types of hyperlipoproteinemias or even familial hypercholesterolemia type B (Bhagavan, 2002; Marcais et al., 2005; Soria et al., 1989), but these are rare and usually more severe in their phenotypes. Confirmation of previously identified associations in different ethnic groups can give additional support to the underlying genetic
architecture of the associated loci, especially when data from related populations are compared (Baba et al., 2009). Genetic structure studies of Europeans have shown that populations from Baltic countries (Estonia, Latvia, and Lithuania), together with Poland and the western part of Russia, form rather a homogeneous group, distinct from the rest of the Europe (Nelis et al., 2009). However, there is little information available on the SNPs associated with blood lipid levels in any of these countries.

Here, we report the associations between common SNPs and the plasma levels of different plasma lipids in a relatively large sample of the Latvian population. The main aims of this study were to investigate the associations between the most-informative SNPs from previous GWAS and four blood lipid parameters: TC, HDL–cholesterol, LDL–cholesterol, and TG in the Latvian population and to provide additional information to characterize the genetic factors that influence blood lipid levels.

Materials and methods

Subjects

We conducted this research using DNA samples from the Genome Database of the Latvian Population (LGDB), which included 18,888 participants in September 2011 when the study sample was selected (Ignatovica et al., 2011). We selected all individuals from this dataset for whom there was information on all four blood lipid parameters (TC, HDL, LDL, and TG), body mass index (BMI), glucose levels, sex, and age, resulting in 1581 samples. We then filtered out subjects with cardiovascular disease and those undergoing lipid-lowering therapies, resulting finally in 1345 samples. One sample was excluded as an outlier because of an extremely high TG level. A proportion (56.5%) of the samples matched those used in a previous study based on the same genotyping panel (Radovica et al., 2013). The genotypes of those samples were obtained from the database, and the remaining 585 samples were genotyped in this study. Written informed consent was acquired from all LGDB participants. The study protocol was approved by the Central Medical Ethics Committee of Latvia (protocol no 2007 A-7 and 01-29.1/25).

SNP data

We previously created a genotyping panel from GWAS, which contained 144 SNPs that were associated with one or more lipid traits (Aulchenko et al., 2009; Burkhardt et al., 2008; Chasman et al., 2008; Edmondson et al.; Heid et al., 2008; Hiura et al., 2009; Kathiresan et al., 2008; Kathiresan et al., 2009; Wallace et al., 2008; Waterworth et al.; Willer et al., 2008; Ma et al.; Pollin et al., 2008; Ridker et al., 2009; Kooner et al., 2008; Sabatti et al., 2009; Sandhu et al., 2008; Saxena et al., 2007; Khovidhunkit et al.). These SNPs occurred in more than 30 loci, including ABCA1, ANGPTL3, APOA1, APOA5, APOB, APOE, CELSR2, CETP, CILP2, DOCK7, FADS2, GALNT2, GCKR, GPR109A/B, HMGCR, LCAT, LDLR, LIPC, LIPG, LPL, MLXIPL, NCAN, NIACR1, NIACR1/KNTC1, NIACR2, NIACR2/GPR81, PCSK9, PLTP, TOMM40, TRIB1, and TTC39B. The SNP selection procedure is described in detail in our previous publication (Radovica et al., 2013).

Genotyping and quality control

All 144 SNPs were genotyped with the Illumina BeadXpress System (Illumina GoldenGate Genotyping Assay), according to the manufacturer’s instructions. The quality control procedure applied to the raw data can be found in our previous article (Radovica et al., 2013). After quality control, the remaining sample consisted of 1273 individuals with 139 genotyped SNPs, with a successful genotyping rate of 99.76%.

Statistical analysis

The normal distributions of all quantitative variables were tested with the two most important parameters, the mean value and its standard deviation (SD) (Altman, 1991), and with the Shapiro–Wilks test. None of the lipid levels were normally distributed according to the Shapiro–Wilks test, even after various transformations. Therefore, to assess the influence of the covariates, we used linear regression, applying less-stringent normality criteria: the 68–95–99.7 rule or the three-sigma rule, according to which about 68% of values should fit within an interval of one SD, 95% in two SDs, and 99.7% in three SDs. Among all the variables tested, the TG levels were...
not normally distributed, so they were log transformed for further statistical analysis. We applied a linear regression analysis with and without the covariates (age, sex, BMI, and glucose levels), and tested epistasis and Hardy–Weinberg equilibrium with the PLINK v2.050 software (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007). The Bonferroni correction was used to calculate the significance level (0.05/139 = 3.5 × 10⁻⁴). To calculate the joint effects, all SNPs in genes with more than one nominally associated SNP were divided into haploblocks using HapMap data and Haplovieview software v4.2 (Barrett, 2009; Barrett et al., 2005), and one SNP was chosen from each haploblock. The joint effect analyses were performed with the SPSS v13.0 software, using a one-sample t test. The association analysis of haplotypes was performed by PLINK toolset. A gene-by-gene interaction analysis was performed with the PLINK v2.050 software and GMDR software Beta 0.9 (http://sourceforge.net/projects/gmdr/)(Lou et al., 2007). Imputation was performed with the IMPUTE2 v2.2.2 software (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#home)(Howie et al., 2009; Marchini et al., 2007) for loci containing more than five SNPs. As reference haplotypes, we used the 1000 Genomes Phase I integrated variant set. The imputation region was set based on the coordinates of the first- and last-tested SNP at each locus. The SNPTEST v2.4.1 software was used to calculate the association between the imputed SNPs and the four lipid parameters (https://mathgen.stats.ox.ac.uk/genetics_software/snpstest/snpstest.html#Download_andCompilation)(Marchini et al., 2007). To accommodate uncertain genotypes, we used the method threshold option, with a threshold value of 0.9. The additive model for association studies was chosen to calculate the P values. The Bonferroni threshold was calculated as follows: 0.05/number of our genotyped SNPs (n = 139) at all loci. Statistical power was calculated with the Quanto v1.2.4 software (Gauderman and Morrison, 2006). The minor allele frequencies (MAFs) of our genotyped SNPs ranged from 0.020 to 0.474. Taking into account this range of MAFs, our study had sufficient power (80%) to detect beta coefficients in the following range for each of the parameters: increased TC, 0.50–0.15; increased LDL–cholesterol, 0.45–0.15; reduced HDL–cholesterol, 0.20; and increased TG, 0.30–0.10.

Results

The clinical characteristics of the study subjects are listed in Table 1.

SNP association tests

We performed linear regression analyses for all 139 SNPs with the four blood lipid parameters (TC, HDL, LDL, and logTG). In unadjusted analyses, 26 SNPs from nine loci were associated with TC, 22 SNPs from 10 loci with LDL, 58 SNPs from 13 loci with HDL, and 40 SNPs from 14 loci with logTG at the nominally significant level (P < 0.05). After age, sex, BMI, and glucose levels were added as covariates to the linear

| Variable | Sample characteristics. |
|----------|-------------------------|
| n        | 1273                    |
| SNPs genotyped | 139                  |
| Genotyping rate, % | 99.76%             |
| Mean age, years ± SD (min–max) | 52.05 ± 13.73 (18–82) |
| Female gender, % | 67.0%                  |
| TC-level, mmol/L ± SD (min–max) | 5.78 ± 1.24 (1.95–11.7) |
| TC, median (Q1: Q3) | 5.70 (4.91; 6.55)     |
| LDL-level, mmol/L ± SD (min–max) | 3.57 ± 1.07 (0.72–8.4) |
| LDL, median (Q1: Q3) | 3.48 (2.81; 4.24)    |
| HDL-level, mmol/L ± SD (min–max) | 1.59 ± 0.44 (0.33–3.52) |
| HDL, median (Q1: Q3) | 1.53 (1.28; 1.87)   |
| TG-level, mmol/L ± SD (min–max) | 1.40 ± 0.75 (0.29–5.87) |
| TG, median (Q1: Q3) | 1.22 (0.89; 1.74)   |
| BMI, kg/m² ± SD (min–max) | 27.02 ± 5.14 (15.64–59.73) |

SD, standard deviation; min, minimum value of variable; max, maximum value of variable; Q1, first quartile of the interquartile range; Q3, third quartile of the interquartile range; TC, total cholesterol; LDL, low-density lipoprotein; HDL, high-density lipoprotein; TG, triglyceride; BMI, body mass index.
regression analysis, 24 SNPs at 11 loci were associated with TC, 23 SNPs at nine loci with LDL, 56 SNPs at 13 loci with HDL, and 40 SNPs at 14 loci with logTG, with the same significance level. Data for the associations of all SNPs are shown in Supplementary Table 1. In total, 16 SNPs remained significantly associated after the Bonferroni correction was applied (Table 2): 10 SNPs at the CETP (cholesteryl ester transfer protein) locus and two SNPs at the MLXIPL (MLX-integrating protein like) locus were associated with reduced HDL–cholesterol levels; one SNP at the TOMM40 (translocase of outer mitochondrial membrane 40) locus was associated with increased LDL–cholesterol; and four SNPs located at the MLXIPL locus were associated with increased logTG levels.

**Haplotype association test**

In order to explore the LD in association with changes in blood lipids detailed haplotype analysis was performed for both genes with more than one significantly associated SNP (MLXIPL and CETP). We thus selected all SNPs from our panel that were located within or in near proximity of CETP and MLXIPL gene and assigned them in haploblocks that were generated from the HapMap1/3 data. In total one haploblock was identified in the MLXIPL locus but four were found in CETP locus. Acquired haplotype data was further used for haplotype association test. Reconstructed haplotypes that exceeded the frequency of 0.01 and the results of association test are shown in Table 3. More detailed description of CETP haplotypes and their genomic localization can be found in our previous paper (Radovica et al., 2013).

**Gene–gene interaction test**

Pair wise SNP–SNP interactions for all 139 SNPs were analyzed with the epistasis option implemented in the PLINK software package to determine whether interactions between two SNPs were associated with lipid levels. We found four pairs of SNPs with \( P \) values <0.0001. One interacting SNP pair was associated with reduced HDL–cholesterol levels (rs1551894 in HMGCR, encoding 3-hydroxy-3-methylglutaryl-CoA reductase, interacting with rs6586891 in LPL, encoding lipoprotein lipase); two SNP pairs were associated with increased LDL–cholesterol levels (rs4803750 in TOMM40 interacting with rs157580 in APOE, encoding apolipoprotein E; and rs157580 in APOE interacting with rs405509 in TOMM40); and one pair of SNPs was associated with increased logTG (rs10889353 in ANGPTL3, encoding angiopotrotein-like 3, interacting with rs166358 in LIPC, encoding hepatic lipase). We also performed a gene–gene epistasis test with the GMDR software, but this analysis revealed no significant gene–gene interactions.

**Imputation**

We performed imputation and subsequent association analyses for 11 loci that contained five or more SNPs (CELSR2, APOB, HMGCR, MLXIPL, LPL, APOA1/5, KNTC1/NIACR1, LIPC, CETP, and TOMM40/APOE), using the 1000 Genomes Phase I integrated variant set of haplotypes as the reference and the coordinates of the outer SNPs from our 11 loci as the interval boundaries. At four loci, APOB (apolipoprotein B), CETP, LPL, and TOMM40, a number of SNPs displayed \( P \) values lower than the Bonferroni threshold for one or more lipid traits. To illustrate these results, we created Manhattan plots for these four loci (shown in Fig. 1). The \( P \) values and \textit{beta} coefficients for the 71 corrected SNPs that were associated with one of the lipid traits after imputation and SNPTEST are shown in Supplementary Table 2. Thus, for APOB and LPL, only the imputed SNPs reached the appropriate significance level for association with LDL, TC, and HDL. One nucleotide deletion at chr2:21428937 (APOB locus) showed the strongest association with LDL cholesterol levels. rs289, located in the sixth intron of the LPL gene, displayed the strongest association with HDL–cholesterol. We found a number of imputed SNPs at the TOMM40 locus that reached the appropriate significance level for their associations with LDL–cholesterol levels, and had lower \( P \) values than our genotyped SNPs. The most strongly associated SNP was rs62117206, located in the intron of the BCL3 gene (B-cell leukemia/lymphoma 3). We found many SNPs with the appropriate significance level at the CETP locus for association with HDL–cholesterol; however none of these SNPs had a \( P \) value lower than that for rs3764261, which was genotyped with our panel. Most of the associated SNPs were found within introns, although some were located in 5’ or 3’ untranslated regions, but none occurred in a coding region.
Table 2
SNPs significantly associated with lipid parameters.

| CHR | SNP            | Corresponding gene | MA  | MAF  | Beta coefficient | P-value | Beta coefficient | P-value | Beta coefficient | P-value | Beta coefficient | P-value | Beta coefficient | P-value after Bonferroni correction |
|-----|----------------|---------------------|-----|------|------------------|---------|------------------|---------|------------------|---------|------------------|---------|------------------|-----------------------------|
| 7   | rs2240466 [6]  | MLXIPL              | T   | 0.109| 0.01839          | 0.80600 | 0.08560          | 0.00048 | -0.04579         | 0.01E-05 | -0.02034         | 0.75340 | -0.01139         | 0.85900                     |
| 7   | rs714052 [6, 16]| MLXIPL              | G   | 0.107| 0.02396          | 0.74670 | 0.08511          | 0.00046 | -0.04716         | 4.76E-05 | -0.03144         | 0.61520 | -0.01319         | 0.85900                     |
| 7   | rs17145738 [15, 21, 26] | MLXIPL          | A   | 0.110| 0.00275          | 0.96970 | 0.08601          | 0.00029 | -0.04625         | 4.44E-05 | -0.01319         | 0.85900 | -0.01319         | 0.85900                     |
| 7   | rs2286276 [6]  | MLXIPL              | T   | 0.154| 0.09463          | 0.08894 | 0.07353          | 5.45E-05 | -0.01940         | 0.02625 | 0.03586          | 0.45600 | -0.01319         | 0.85900                     |
| 7   | rs11974409 [6, 9]| MLXIPL            | C   | 0.164| 0.01688          | 0.78570 | 0.07183          | 0.00041 | -0.03482         | 0.00034 | -0.00917         | 0.86430 | -0.01319         | 0.85900                     |
| 16  | rs173539 [11, 16]| CETP               | T   | 0.337| 0.08240          | 0.10270 | 0.09744          | 3.39E-09 | -0.01106         | 0.16330 | 0.02580          | 0.55440 | 0.03005          | 0.49460                     |
| 16  | rs3764261 [6, 9, 11, 21, 26] | CETP               | T   | 0.329| 0.08882          | 0.08094 | 0.10050          | 1.46E-09 | -0.01126         | 0.15850 | 0.01733          | 0.66070 | 0.03005          | 0.49460                     |
| 16  | rs1800775 [6, 11, 15, 22, 24, 26, CETP | C | 0.118| -0.05645       | 0.21770 | -0.08538        | 1.02E-08 | 0.00723         | 0.31570 | 0.01733          | 0.66070 | 0.01733          | 0.66070                     |
| 16  | rs711752 [11, 25] | CETP               | T   | 0.456| 0.06145          | 0.20260 | 0.08489          | 7.19E-08 | -0.00520         | 0.49240 | -0.01202         | 0.77320 | -0.01202         | 0.77320                     |
| 16  | rs708272 [24]  | CETP               | A   | 0.455| 0.05463          | 0.25800 | 0.08567          | 5.59E-08 | -0.00541         | 0.47550 | -0.01941         | 0.64200 | -0.01941         | 0.64200                     |
| 16  | rs9929488 [11] | CETP               | G   | 0.278| -0.04669         | 0.37040 | -0.07313         | 1.85E-05 | 0.01412          | 0.08442 | 0.00417          | 0.92640 | 0.00417          | 0.92640                     |
| 16  | rs7203984 [6, 11] | CETP              | G   | 0.195| -0.00078         | 0.98930 | -0.08570         | 6.23E-06 | 0.02179          | 0.01649 | 0.05610          | 0.26290 | 0.05610          | 0.26290                     |
| 16  | rs9939924 [11] | CETP               | T   | 0.194| 0.00023          | 0.99690 | -0.09567         | 8.00E-07 | 0.02311          | 0.01123 | 0.06698          | 0.18210 | 0.06698          | 0.18210                     |
| 16  | rs7205804 [6, 11, 25] | CETP              | T   | 0.457| 0.02881          | 0.54900 | 0.07848          | 5.69E-07 | -0.00807         | 0.28510 | -0.03287         | 0.42900 | -0.03287         | 0.42900                     |
| 16  | rs11076715 [11] | CETP               | G   | 0.178| 0.00222          | 0.90760 | -0.0815          | 7.84E-06 | 0.02283          | 0.01570 | 0.05980          | 0.25090 | 0.05980          | 0.25090                     |
| 19  | rs4803750 [6, 9, 27] | TOMM40           | C   | 0.080| -0.23300         | 0.00634 | 0.02782          | 0.12170 | 0.03400          | 0.01119 | -0.3820          | 2.85E-05 | -0.3820          | 2.85E-05                     |

SNPs with $P$ values $<0.05$ are marked in bold; SNPs with $P$ values $<0.05$ after Bonferroni correction are black highlighted with white letters. All calculations were made with a linear regression analysis with covariates (age, sex, BMI, and glucose levels). CHR, chromosome; SNP, single-nucleotide polymorphism; MA, minor allele; MAF, minor allele frequency; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

I. Radovica et al. / Meta Gene 2 (2014) 565–578
### Table 3

Haplotypes in CETP and MLXIPL genes associated with blood lipid levels.

| Gene | ID | Haplotype sequence | FRQ | TC Beta | TC P-value | HDL Beta | HDL P-value | LDL Beta | LDL P-value | LogTG Beta | LogTG P-value |
|------|----|-------------------|-----|--------|-----------|----------|------------|----------|------------|------------|--------------|
| CETP | 1.1 | G C               | 0.084 | 0.05723 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 1.2 | A T               | 0.916 | -0.05723 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 2.1 | G T T T          | 0.326 | 0.08774 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 2.2 | C T G            | 0.444 | -0.08692 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 2.3 | A C C G          | 0.134 | 0.03950 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 2.4 | G C C G         | 0.028 | -0.11970 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 2.5 | A C T G         | 0.057 | 0.03455 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 3.1 | A T A G C T G T A | 0.442 | 0.02923 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 3.2 | C C G G T C G   | 0.171 | 0.02353 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 3.3 | C C G G C T T C A | 0.015 | -0.01085 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 3.4 | A C G A C T G C A | 0.028 | 0.29760 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 3.5 | C C G G C T G C A | 0.186 | -0.07357 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 3.6 | A C G G G T G C A | 0.044 | -0.08910 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 3.7 | C C G A C T G C A | 0.028 | -0.11310 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 3.8 | C C G G G T G C A | 0.036 | -0.09055 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 4.1 | C C               | 0.971 | -0.15460 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 4.2 | C T               | 0.020 | 0.06946 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
| MLXIPL | 5.1 | T C G A T C      | 0.106 | -0.00579 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 5.4 | C T G A G C T    | 0.770 | -0.09202 | -0.03150 | -0.04205 | 0.01601 | 0.27640 |
|      | 5.2 | C C A G T T      | 0.054 | 0.00051 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |
|      | 5.3 | C T A G T T      | 0.065 | 0.24710 | 0.03150 | 0.04205 | 0.01601 | 0.27640 |

FRQ, frequency of haplotypes; TC, total cholesterol; HDL, high density lipoprotein cholesterol; LDL, low density lipoprotein cholesterol; logTG, logarithmic value of triglyceride level; ID, haplotype identification number (first digit—number of haploblock studied; second digit—number of individual haplotype); first haploblock of CETP gene consists of rs2241770 and rs16962767; second haploblock of CETP gene consists of rs12448528, rs173539, rs12708967 and rs3764261; third haploblock of CETP gene consists of rs1800775, rs711752, rs708272, rs1864163, rs929488, rs7203984, rs9939224, rs7205804 and rs1176175; fourth haploblock of CETP gene consists of rs5880 and rs1800777; first haploblock of MLXIPL gene consists of rs2240466, rs1178979, rs714052, rs17145738, rs2286276 and rs11974409; haplotypes considered as “protective” are highlighted in grey; haplotypes considered as “risk” are highlighted in black with white letters; P-values lower than 0.05 are highlighted in black with white letters. 

I. Radovica et al. / Meta Gene 2 (2014) 565–578
Joint effect analysis

Because the effects of individual SNPs are rather small, we decided to investigate the summary effects of associated SNPs on blood lipid parameters, using the risk allele dosage test and those SNPs that were nominally associated with different lipid levels before the Bonferroni correction, each of which represented one haplобрlock. Sixteen SNPs representing the ABCA1, ANGPTL3, APOB, APOE, Dock7, HMGCR, LCAT, LIPC, and TOMM40 loci were included in case of TC; 24 SNPs representing the ABCA1, APOA1, CETP, HMGCR, NIACR1, LCAT, LIPC, LPL, MLXIPL, PLTP, and TRIB1 loci were included in cases of HDL cholesterol; 15 SNPs representing ABCA1, ANGPTL3, APOA1, APOB, APOE, Dock7, HMGCR, LCAT, and TOMM40 loci were included in case of LDL–cholesterol; and 27 SNPs representing ABCA1, ANGPTL3, APOA1, APOA5, APOB, CETP, CILP2, Dock7, GCKR, LIPC, LPL, MLXIPL, NCA7, and TOMM40 loci were included in case of logTG. The distribution of risk allele number within each group of lipid parameters analyzed was almost normally distributed and therefore we divided the sample into three equivalent groups based on the range of the risk allele distribution separately for each lipid parameter. We then performed a one-sample t-test with the SPSS software to calculate the differences between these three groups. We found that the number of risk alleles correlated strongly with the blood lipid levels for all four parameters. The results of this analysis are shown in Fig. 2. In case of TC the group of individuals with 4–10 risk alleles had a mean TC level 0.53 mmol/L lower than that of the median group, represented by individuals with 11–18 risk alleles (P = 1.59E – 44). In the group with 19–26 risk alleles, the mean TC level was 0.40 mmol/L higher (P = 4.18E – 97) than that in the median group. The mean HDL–cholesterol level was 0.24 mmol/L higher (P = 1.59E – 44) in the group with 13–21 risk alleles and 0.18 mmol/L lower (P = 7.51E – 241) in the group with 31 – 40 risk alleles compared to the value for the median group (22–30 risk alleles). The mean LDL–cholesterol level was 0.56 mmol/L lower (P = 1.36E – 22) in the group with two to nine risk alleles and 0.51 mmol/L higher (P = 7.91E – 140) in the group with 10–17 risk alleles compared to the value in the median group (10–17 risk alleles). The mean logTG value was 0.07 log mmol/L lower (P = 1.06E – 31) in the group with 13–21 risk alleles and 0.09 log mmol/L higher (P = 1.76E – 34) in the group with 30–39 risk alleles compared to the median group (22–29 risk alleles).

Discussion

The aim of this study was to identify associations between 144 SNPs selected from published GWAS and blood lipid levels in the Latvian population. This is the first report in which many genetic loci, involved in a number of lipid-related metabolic pathways, were simultaneously studied in a relatively large group representing the Latvian population (420 males, 853 females; for detailed data, see Table 1). Our study has convincingly identified associations between 16 of the 139 most strongly associated SNPs and one or more blood lipid parameters.

In our previous case–control study, we reported strong associations between 10 SNPs in CETP and reduced HDL–cholesterol (Radovica et al., 2013). In this study, we confirmed this strong evidence for this association between the CETP gene and HDL–cholesterol levels and discovered the new association of two SNPs at the MLXIPL locus (Table 2), using quantitative analyses. Moreover, haplotype analyses also confirmed the associations demonstrated in our previous study (Radovica et al., 2013). These results also agree well with those of other studies around the world (Chasman et al., 2008; Heid et al., 2008; Hiura et al., 2009; Kathiresan et al., 2008; Kathiresan et al., 2009; Willer et al., 2008; Ma et al.; Ridker et al., 2009; Khovidhunkit et al.; Thompson et al., 2005; Thompson et al., 2007) and are supported by the well-known function of CETP in lipid metabolism. CETP promotes the transfer of cholesteryl esters from HDL particles to apolipoprotein-B-containing particles in exchange for TG, allowing the receptor-mediated uptake of cholesterol esters by the liver and reducing HDL–cholesterol levels (Vance and Vance, 2008). In contrast, CETP deficiency increases HDL–cholesterol levels. Supporting the association between SNPs in the MLXIPL locus and HDL–cholesterol, several reports have linked MLXIPL and HDL–cholesterol. However, their relationship within the lipid metabolic pathways is not as direct as that of the SNP–CETP association. The
protein encoded by the MLXIPL gene is a helix–loop–helix leucine zipper transcription factor, which forms a heterodimer with MLX to bind DNA (Cairo et al., 2001). This transcription factor complex is activated by high glucose and inhibited by cyclic adenosine monophosphate (cAMP) (Kawaguchi et al., 2001), and can repress E-box-dependent transcription (Cairo et al., 2001) or the carbohydrate response element motifs in the promoters of lipogenic enzymes. In adipose tissues, MLXIPL is a major determinant of adipose tissue fatty-acid synthesis and systemic insulin sensitivity (Herman et al.). Therefore, it is plausible that changes in MLXIPL expression indirectly influence HDL–cholesterol levels through various metabolic pathways.

We also found strong associations between four SNPs at the MLXIPL locus and increased logTG (Table 2). Because MLXIPL is a transcription factor that influences fatty-acid synthesis in adipose tissue, it is plausible that SNPs in this gene indirectly affect TG levels. We also found that a SNP in the TOMM40 locus is associated with increased LDL–cholesterol levels (Table 2). More precisely, this SNP is located in the first intron of the BCL3 gene (Fig. 1 TOMM40 locus). The protein encoded by the BCL3 gene is an inhibitor of subunit two of nuclear factor kappa-B (NF-κB) (Franzoso et al., 1992; Wulczyn et al., 1992). Because interactions between BCL3 and NF-κB have not yet been linked to lipid metabolism, we think that this SNP might be in linkage disequilibrium (LD) with functional SNPs in one of the genes encoding various apolipoproteins located downstream from the TOMM40 gene. For example, mutations in the APOE gene (located immediately downstream from TOMM40) can cause hyperlipoproteinemia type III (Utermann et al., 1979). Alternatively, the associated SNPs may be located within upstream regulatory elements of these apolipoprotein genes because the distance between the TOMM40 gene and, for example, the APOE gene is rather large (2.1 Mbp), arguing against a simple LD-based explanation.

Since the number of associated SNPs was found in CETP and MLXIPL gene, we estimated the association of individual haplotypes with all lipid levels. First it should be noted that in case of CETP the results of HDL associations corresponded exactly to our previous findings (Radovica et al., 2013). Most of the associated SNPs fall in second and third haploblock (Table 3). These haploblocks cover the region starting from ~2 kb upstream CETP gene, promoter region and include first ~10 kb of CETP gene. Interestingly, the strongest association was observed for two protective haplotypes with increased HDL levels: haplotype 2.1 from second haploblock and 3.1 from the third haploblock. For second haploblock it is clear that effect is explained by the presence of risk alleles from two SNPs rs173539 and rs3764261 that are both in strong LD with each other haploblock and 3.1 from the third haploblock. For second haploblock it is clear that effect is explained by the presence of risk alleles from two SNPs rs173539 and rs3764261 that are both in strong LD with each other haploblock and 3.1 from the third haploblock. For second haploblock it is clear that effect is explained by the presence of risk alleles from two SNPs rs173539 and rs3764261 that are both in strong LD with each other haploblock and 3.1 from the third haploblock.
which is synthesized in the liver and acts in the uptake of lipoproteins and cholesteryl esters by hepatocytes, thus directly influencing plasma lipid levels (Cai et al., 1989). It is not known how these two proteins interact to influence lipid metabolism, but such interactions cannot be excluded. However, it should be noted that the majority of genes selected for this study have well-established functions in lipid metabolism. Therefore, all attempts to explain these interactions statistically, based on the relationships of these protein functions, are highly biased, and in all cases, further investigation is required in large well-characterized cohorts. Unfortunately, these findings were not supported by the GMDR software, perhaps because more-specific algorithms are used in GHMDR, taking into account, for example, covariates, “course of dimensionality”, and other criteria that are not included in the PLINK toolset and which therefore better exclude false positive results (Chen et al.).

For a more detailed analysis and to find other strongly associated SNPs at our test loci, we performed imputation for 11 loci that contained five or more SNPs (CELSR2, APOB, HMGR, MLXIPL, LPL, APOA1/5, KNTC1/NIACR1, LIPC, CETP, and TOMM40/APOE). We found many imputed SNPs in the CETP gene that were associated with changes in HDL–cholesterol levels; although none of these associations was stronger than those of the originally genotyped SNPs (Fig. 1 CETP locus). These results again agree well with those of our previous study, in that CETP is the strongest and most common genetic factor to influence HDL–cholesterol levels in the Latvian population. Imputed SNPs at another three loci, APOB, LPL, and TOMM40, were significantly associated with one or more lipid trait (Fig. 1). All these SNPs were located in noncoding regions and many of them have already been associated with disease phenotypes, including CAD, myocardial infarction, and atherosclerosis. The APOB and LPL genes are known to be directly involved in lipid metabolism, and a number of functional mutations in these genes have already been reported to cause different types of monogenic familial hypercholesterolemia (http://www.hgmd.org/). Therefore, it is likely that common variants with smaller effects may play roles in milder forms of dyslipidemia. It is also possible that our cohort contained a number of patients with undiagnosed monogenic hypercholesterolemia. This may have influenced the association results either because of the LD between these mutations and the markers used in our study or simply by adding a stronger genetic modifier. A number of imputed SNPs at the TOMM40 locus reached the appropriate significance level in their associations with LDL–cholesterol levels. Again, the role of TOMM40 in the regulation of LDL metabolism is unclear (see discussion above).

To achieve our aim, we performed an allele dosage association test with nominally associated SNPs, each representing one haploblock, for all four blood lipid parameters to calculate the influence of the number of risk alleles on lipid levels. We divided our sample in three equivalent groups based on the number of risk alleles. There was a strong correlation between the number of risk alleles and the plasma lipid levels for all four lipid parameters (Fig. 2). Therefore, it is possible that the discovery and inclusion of additional lipid-influencing SNPs in similar analyses will lead to the development of prognostic tests for dyslipidemia and the prescription of appropriate drugs.

Conclusions

The results of this study confirm that of all the genes included in the analysis, CETP, MLXIPL, and TOMM40 are the genetic factors that most strongly influence the plasma lipid levels in our study population. We have also shown that these risk alleles have strong cumulative effects on all lipid parameters. It should be noted that this is the first replication study of a relatively large sample population from the Baltic States and Eastern Europe.

Abbreviations

| Abbreviation | Description                      |
|--------------|----------------------------------|
| CAD          | coronary artery disease          |
| TC           | total cholesterol                |
| SNP          | single-nucleotide polymorphism   |
| LDL          | low-density lipoprotein          |
| HDL          | high-density lipoprotein         |
| TG           | triglyceride                     |
| BMI          | body mass index                  |
| MAF          | minor allele frequency           |
| CHR          | chromosome                       |
| MA           | minor allele                     |
| GWAS         | genome-wide association studies  |
cAMP  cyclic adenosine monophosphate  
LD  linkage disequilibrium  
SD  standard deviation

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2014.07.006.

Acknowledgments

The authors thank the Genome Database of the Latvian Population and the Latvian Biomedical Research and Study Centre for providing data and DNA samples.

This work was supported by the Latvian Research Program "Biomedicine for health" and the European Social Fund within the project "Support for Doctoral Studies at the University of Latvia", contract no. 2009/0138/1DP/1.1.2.1.2/09/IPIA/VIAA/004.

References

Altman, D.G., 1991. Practical Statistics for Medical Research. In: Altman, D.G. (Ed.), Chapman & Hall/CRC, London.
Aulchenko, Y.S., et al., 2009. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat. Genet. 41 (1), 47–55.
Baba, K., et al., 2009. Validating an alternate version of the chewing function questionnaire in partially dentate patients. BMC Oral Health 9, 9.
Barrett, J.C., 2009. Haploview: visualization and analysis of SNP genotype data. Cold Spring Harb Protoc 2009 (10) (p. pdb ip71).
Braun, T.R. et al., 2008. A replication study of GWAS-derived lipid genes in Asian Indians: the chromosomal region 11q23.3 harbors loci contributing to triglycerides. PLoS One. 7(5), e37056.
Burkhardt, R., et al., 2008. Common SNPs in HMGCR in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. Arterioscler. Thromb. Vasc. Biol. 28 (11), 2078–2084.
Cai, S.J., et al., 1989. Structure of the human hepatic triglyceride lipase gene. Biochemistry 28 (23), 8966–8971.
Caire, S., et al., 2001. WBSCR14, a gene mapping to the Williams–Beuren syndrome deleted region, is a new member of the Mix transcription factor network. Hum. Mol. Genet. 10 (6), 617–627.
Camenisch, G., et al., 2002. ANGPTL3 stimulates endothelial cell adhesion and migration via integrin alpha vbeta 3 and induces blood vessel formation in vivo. J. Biol. Chem. 277 (19), 17281–17290.
Chasman, D.I., et al., 2008. Genetic loci associated with plasma concentration of low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, apolipoprotein A1, and apolipoprotein B among 6382 white women in genome-wide analysis with replication. Circ. Cardiovasc. Genet. 1 (1), 21–30.
Chen, G.B. et al., Practical and theoretical considerations in study design for detecting gene-gene interactions using MDR and GMDR approaches. PLoS One. 6(2), e16981.
Dachet, C., et al., 2000. New functional promoter polymorphism, CETP/-629, in cholesteryl ester transfer protein (CETP) gene related to CETP mass and high density lipoprotein cholesterol levels: role of Sp1/Sp3 in transcriptional regulation. Arterioscler. Thromb. Vasc. Biol. 20 (2), 507–515.
Durrington, P., 2003. Dyslipidaemia. Lancet 362 (9385), 717–731.
Edmondson, A.C. et al., Dense genotyping of candidate gene loci identifies variants associated with high-density lipoprotein cholesterol. Circ. Cardiovasc. Genet. 4(2), 145–155.
Franzoso, G., et al., 1992. The candidate oncprotein Bcl-3 is an antagonist of p50/NF-kappa B-mediated inhibition. Nature 359 (6393), 339–342.
Gauderman, W. & Morrison, J., 2006. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. http://hydra.usc.edu/gxe.
Goldberg, A.C. Dyslipidemia. 2013, [cited 2013 31.01.2014]; Available from: http://www.merckmanuals.com/professional/endocrine_and_metabolic_disorders/lipid_disorders/dyslipidemia.html.
Heid, I.M., et al., 2008. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. Circ. Cardiovasc. Genet. 1 (1), 10–20.
Herman, M.A. et al., A novel ChREBP isoform in adipose tissue regulates systemic glucose metabolism. Nature. 484(7394), 333–338.
Hiura, Y., et al., 2009. Identification of genetic markers associated with high-density lipoprotein-cholesterol by genome-wide screening in a Japanese population: the Suita study. Circ. J. 73 (6), 1119–1126.
Howie, B.N., Donnelly, P., Marchini, J., 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 5 (6), e1000529.
Igl, W., et al., Modeling of environmental effects in genome-wide association studies identifies SLC2A2 and HP as novel loci influencing serum cholesterol levels. PLoS Genet. 6(1), e1000798.
Ignatovic, V., et al., 2011. Single nucleotide polymorphisms of the purinergic 1 receptor are not associated with myocardial infarction in a Latvian population. Mol. Biol. Rep. 39 (2), 1917–1925.
Johansen, C.T. et al., Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat. Genet. 42(8), 684–687.
Kathiresan, S., et al., 2008. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat. Genet. 40 (2), 189–197.

Kathiresan, S., et al., 2009. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat. Genet. 41 (1), 56–65.

Kawaguchi, T., et al., 2001. Glucose and cAMP regulate the L-type pyruvate kinase gene by phosphorylation/dephosphorylation of the carbohydrate response element binding protein. Proc. Natl. Acad. Sci. U. S. A. 98 (24), 13710–13715.

Khovidhunkit, W., et al., Resequencing CETP, LIPC and LIPG Genes in Thai Subjects With Hyperalphalipoproteinemia. Am. J. Cardiol. 62, 149–151.

Lanktree, M.B., et al., 2009. Replication of genetic associations with plasma lipoprotein traits in a multiethnic sample. J. Lipid Res. 50 (7), 1487–1496.

Lou, X.Y., et al., 2007. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. Am. J. Hum. Genet. 80 (6), 1125–1137.

Ma, L., et al., Genome-wide association analysis of total cholesterol and high-density lipoprotein cholesterol levels using the Framingham heart study data. BMC Med. Genet. 11, 55.

Marchini, J., et al., 2007. A new multipoint method for genome–wide association studies by imputation of genotypes. Nat. Genet. 39 (7), 906–913.

Marchini, J., et al., 2005. Exome sequencing, ANGPTL3 mutations, and familial combined hyperlipidemia. N. Engl. J. Med. 363 (23), 2220–2227.

Nelis, M., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4 (5), e5472.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Marchini, J., et al., 2007. A new multipoint method for genome–wide association studies by imputation of genotypes. Nat. Genet. 39 (7), 906–913.

Marchini, J., et al., 2005. Exome sequencing, ANGPTL3 mutations, and familial combined hyperlipidemia. N. Engl. J. Med. 363 (23), 2220–2227.

Nelis, M., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4 (5), e5472.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Marchini, J., et al., 2007. A new multipoint method for genome–wide association studies by imputation of genotypes. Nat. Genet. 39 (7), 906–913.

Marchini, J., et al., 2005. Exome sequencing, ANGPTL3 mutations, and familial combined hyperlipidemia. N. Engl. J. Med. 363 (23), 2220–2227.

Nelis, M., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4 (5), e5472.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Marchini, J., et al., 2007. A new multipoint method for genome–wide association studies by imputation of genotypes. Nat. Genet. 39 (7), 906–913.

Marchini, J., et al., 2005. Exome sequencing, ANGPTL3 mutations, and familial combined hyperlipidemia. N. Engl. J. Med. 363 (23), 2220–2227.

Nelis, M., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4 (5), e5472.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Marchini, J., et al., 2007. A new multipoint method for genome–wide association studies by imputation of genotypes. Nat. Genet. 39 (7), 906–913.

Marchini, J., et al., 2005. Exome sequencing, ANGPTL3 mutations, and familial combined hyperlipidemia. N. Engl. J. Med. 363 (23), 2220–2227.

Nelis, M., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4 (5), e5472.

Pollin, T.I., et al., 2008. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 322 (5908), 1702–1705.

Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.