

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Surakka, Ida; Whitfield, JB; Perola, M; Visscher, PM; Montgomery, GW; Falchi, M; Willemsen, G; Geus, EJC de; Magnusson, PKE; Christensen, K; Sorensen, TIA; Pietiläinen, KH; Rantanen, Taina; Silander, K; Widén, E; Muilu, J; Rahman, I; Liljedahl, U; Syvänen, A-C; Palotie, A; Kaprio, J; Kyvik, KO; Pedersen, NL; Boomsma, DI; Spector, T; Martin, NG; Ripatti, S; Peltonen, L

**Title:** A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol

**Year:** 2012

**Version:** Published version

**Copyright:** © The Authors 2012

**Rights:** In Copyright

**Rights url:** <http://rightsstatements.org/page/InC/1.0/?language=en>

**Please cite the original version:**

Surakka, I., Whitfield, J., Perola, M., Visscher, P., Montgomery, G., Falchi, M., Willemsen, G., Geus, E. D., Magnusson, P., Christensen, K., Sorensen, T., Pietiläinen, K., Rantanen, T., Silander, K., Widén, E., Muilu, J., Rahman, I., Liljedahl, U., Syvänen, A.-C., . . . Peltonen, L. (2012). A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. *Twin Research and Human Genetics*, 15(6), 691-699. <https://doi.org/10.1017/thg.2012.63>

# A Genome-Wide Association Study of Monozygotic Twin-Pairs Suggests a Locus Related to Variability of Serum High-Density Lipoprotein Cholesterol

Ida Surakka,<sup>1,2</sup> John B. Whitfield,<sup>3</sup> Markus Perola,<sup>1,2,4</sup> Peter M. Visscher,<sup>5</sup> Grant W. Montgomery,<sup>3</sup> Mario Falchi,<sup>6</sup> Gonneke Willemsen,<sup>7</sup> Eco J. C. de Geus,<sup>7</sup> Patrik K. E. Magnusson,<sup>8</sup> Kaare Christensen,<sup>9,10</sup> Thorkild I. A. Sørensen,<sup>11,23</sup> Kirsi H. Pietiläinen,<sup>1,12</sup> Taina Rantanen,<sup>13</sup> Kaisa Silander,<sup>1</sup> Elisabeth Widén,<sup>1</sup> Juha Muiilu,<sup>1</sup> Iffat Rahman,<sup>8</sup> Ulrika Liljedahl,<sup>14</sup> Ann-Christine Syvänen,<sup>14</sup> Aarno Palotie,<sup>1,15,16,17</sup> Jaakko Kaprio,<sup>1,18,19</sup> Kirsten O. Kyvik,<sup>20,21</sup> Nancy L. Pedersen,<sup>8</sup> Dorret I. Boomsma,<sup>7</sup> Tim Spector,<sup>22</sup> Nicholas G. Martin,<sup>3</sup> Samuli Ripatti,<sup>1,2,15</sup> and Leena Peltonen\* on behalf of the GenomEUtwin Project

<sup>1</sup>FIMM, Institute for Molecular Medicine, Finland Biomedicum FI-00014, University of Helsinki, Finland

<sup>2</sup>Public Health Genomics Unit, National Institute for Health and Welfare, FI-00271 Helsinki, Finland

<sup>3</sup>Queensland Institute of Medical Research, Locked Bag 2000, Royal Brisbane Hospital, Herston, Queensland, Australia

<sup>4</sup>The Estonian Genome Center and the Center of Translational Genomics of the University of Tartu, Tartu, Estonia

<sup>5</sup>Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia

<sup>6</sup>Genomic Medicine, Hammersmith Hospital, Imperial College London, London, UK

<sup>7</sup>Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands

<sup>8</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

<sup>9</sup>The Danish Twin Registry, University of Southern Denmark, and Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark

<sup>10</sup>Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense C, Denmark

<sup>11</sup>Institute of Preventive Medicine, Copenhagen University Hospital, Copenhagen, Denmark

<sup>12</sup>Obesity Research Unit, Department of Medicine, Division of Endocrinology, Helsinki University Central Hospital and University of Helsinki, Finland

<sup>13</sup>Department of Health Sciences, Gerontology Research Centre, University of Jyväskylä, Finland

<sup>14</sup>Department of Medical Sciences, Uppsala University, Sweden

<sup>15</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

<sup>16</sup>Department of Medical Genetics, Haartman Institute, University of Helsinki and Helsinki University Central Hospital, Finland

<sup>17</sup>The Broad Institute, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>18</sup>Department of Public Health, Faculty of Medicine, FI-00014 University of Helsinki, Finland

<sup>19</sup>Unit for Child and Adolescent Mental Health, National Institute for Health and Welfare, Helsinki, Finland

<sup>20</sup>Institute of Regional Health Services Research, University of Southern Denmark, Odense C, Denmark

<sup>21</sup>Odense Patient Data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark

<sup>22</sup>Department of Twin Research and Genetic Epidemiology, King's College London, London, UK

<sup>23</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen

Genome-wide association analysis on monozygotic twin-pairs offers a route to discovery of gene-environment interactions through testing for variability loci associated with sensitivity to individual environment/lifestyle. We present a genome-wide scan of loci associated with intra-pair differences in serum lipid and apolipoprotein levels. We report data for 1,720 monozygotic female twin-pairs from GenomEUtwin project with 2.5 million SNPs, imputed or genotyped, and measured serum lipid fractions for both twins. We found one locus associated with intra-pair differences in high-density lipoprotein cholesterol, *rs2483058* in an intron of *SRGAP2*, where twins carrying the C allele are more sensitive to environmental factors ( $P=3.98 \times 10^{-8}$ ). We followed up the association in further genotyped monozygotic twins ( $N=1,261$ ), which showed a moderate association for the variant ( $P=0.200$ , same direction of an effect). In addition, we report a new association on the level of apolipoprotein A-II ( $P=4.03 \times 10^{-8}$ ).

■ **Keywords:** twins, association, lipids, apolipoproteins, interaction

RECEIVED 31 August 2012; ACCEPTED 3 September 2012. First published online 3 October 2012.

\*Deceased.

ADDRESS FOR CORRESPONDENCE: Samuli Ripatti, Institute for Molecular Medicine, Finland, FIMM, PO Box 20, University of Helsinki, Finland. E-mail: samuli.ripatti@fimm.fi

Serum lipid levels have been associated with cardiovascular diseases, metabolic syndrome, and type II diabetes (Kannel et al., 1961; Miller & Miller, 1975; Pilia et al., 2006). Variation in lipids levels is highly influenced by heritable factors (Friedlander et al., 1997), and 95 loci have already been associated with levels of high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides (TG), and total cholesterol (TC) in numerous study samples and replicated in various populations using genome-wide approaches (Aulchenko et al., 2008; Kathiresan et al., 2008; Kooner et al., 2008; Teslovich et al., 2010; Willer et al., 2008). However, the genetic associations affecting blood apolipoproteins have been studied mainly with the candidate gene approach.

Data emerging from epidemiological studies imply that genes interact with environmental factors causing additional variability in lipid levels (Thorn et al., 1998; Visscher & Duffy, 2006). In animal and plant models,  $G \times E$  interactions can be studied by controlling the environmental factors. In human populations, the environment of an individual is highly variable and the individual can choose environments. This makes it difficult to find loci interacting with specified environmental factors. Recently published screen for  $G \times E$  interactions identified a locus where, depending on the genotype at that locus, the connection between waist-to-hip ratio and total cholesterol differed (Surakka et al., 2011), but to detect the association, over 40,000 individuals from various population cohorts were used.

The total environmental sensitivity of a locus, without specifying which environmental factors are salient, can be tested in general population samples by comparing the trait variances between different genotype classes in genome-wide association (GWA) studies. For this sensitivity association screening, monozygotic (MZ) twins offer the additional benefit of controlling for possible confounding due to possible epistatic effects. MZ twin-pairs share essentially identical genomes at the sequence level and the differences within pairs; for example, lipid levels reflect environmentally influenced differences. Genetic variation may determine the extent to which environmental differences lead to phenotypic within pair variability (Magnus et al., 1981). An association with within-MZ-pair trait differences directly indicates gene–environment interaction, where one allele increases or restricts the effect of environmental factors or environmental factors affect the gene expression differently for different alleles. This is seen as elevated or diminished levels of variability in trait values. In a recent methodological paper, the power to detect variability association using either difference between twins or variability in population approaches was evaluated (Visscher & Posthuma, 2010). The study showed that using MZ twins is more powerful than the approach using population variances, when the trait under examination has high within-pair correlations in MZ twin-pairs.

In this study we have screened for genetic “variability” loci where the loci are associated with differences in lipid levels between the two members of a pair in our GWA study in MZ female twins. More specifically, the aim of our study was to search variability genes for HDL, LDL, TC, TG, and apolipoprotein A-I, A-II, B, and E (APOA1, APOA2, APOB, and APOE respectively). For the apolipoprotein data at hand we have additionally screened for apolipoprotein level associations.

## Materials and Methods

### Study Samples

The Danish, Dutch, Finnish, Italian, Norwegian, and Swedish national twin cohorts, together with St Thomas’ twin cohort in the United Kingdom and an Australian twin cohort, form the basis of the GenomEUtwin project ([www.genomeutwin.org](http://www.genomeutwin.org)), a collection of 800,000 twins (Peltonen & GenomEUtwin, 2003). From these cohorts, female MZ pairs, aged 20–80 years, with blood lipid measurements available on both members were selected, and the sample consisted of pairs from Australia (459 pairs), Denmark (173 pairs), Finland (152 pairs), the Netherlands (331 pairs), Sweden (301 pairs), and the United Kingdom (462 pairs). DNA samples from one member of each pair were used for genotyping.

### Verification of Zygosity

To verify the zygosity and monitor for possible sample mix-ups, a subset of the samples were subjected to genotyping of microsatellite or single nucleotide polymorphism (SNP) panels prior to the genome-wide analysis. The zygosity of all twin-pairs from the United Kingdom and Finland was verified by genotyping both twins of each twin-pair against a marker panel assaying of nine polymorphic autosomal (AC00818, D1S425, D2S338, D10S1651, D13S263, D16S539, D18S1102, D20S448, and GDB371492) and one X-linked microsatellite loci (DXS1227) in addition to the amelogenin locus yielding X- and Y-chromosomal-specific bands. The zygosity of the Danish samples was primarily assessed by nine DNA-based microsatellite markers with the PE Applied Biosystems AmpFISTR Profiler Plus Kit. For 10 twin-pairs, the zygosity was verified using the same microsatellite panel as described above. The zygosity of the Australian twin-pairs was verified using a commercial kit (AmpFISTR Profiler Plus Amplification KIT, Applied Biosystems, Foster City, CA) that analyzed nine independent, highly polymorphic DNA markers, plus the amelogenin marker for sex (Nyholt, 2006). The zygosity of the Swedish twin-pairs was verified by genotyping of a validated panel of 47 SNPs as described in Hannelius et al. (2007).

### Genotyping and Quality Control

The samples were genotyped using the Infinium II assay (Steemers et al., 2006) on the HumanHap300-Duo

Genotyping BeadChips (Illumina Inc, San Diego, USA). Genotyping was performed according to the manufacturers' instructions (Infinium II Assay Two-Sample Manual, #11230506 rev. A, Illumina Inc.). In total, 318,237 SNPs were genotyped on the BeadChips. The signal intensity data were converted into genotypes using the Illumina Beadstudio 2.0 software. Genotypes were assigned using custom-defined cluster positions generated from samples genotyped in-house. A total of 1,922 subjects were genotyped, of which two-thirds were processed and genotyped at the Finnish Genome Center in Helsinki and one-third at the SNP Technology Platform, Uppsala University, Uppsala, Sweden ([www.genotyping.se](http://www.genotyping.se)). To verify the consistency of genotyping between laboratories, 14 duplicate samples were genotyped at both sites. The generated ~4.4 million genotypes were compared for consistency, yielding a reproducibility rate of 99.99%. The quality of samples and reagents was monitored by sample call rates, sample heterozygosity rates, and sex. Twenty-eight samples yielding <95% SNP call rates were discarded. Altogether, successful genotyping was performed on 1,868 unrelated subjects, of which 93.5% had a call rate of >99%. The average sample heterozygosity rate varied between 30.9% and 39.1%.

### Imputation

Before imputation, individuals of non-European origin and with genotyping frequency of lower than 95% were excluded. These resulted in 1,720 individuals (Australia = 421, Denmark = 168, Finland = 137, the Netherlands = 275, Sweden = 297, and the United Kingdom = 422) in the final analysis set. SNPs having minor allele frequency of lower than 1%, Hardy Weinberg Equilibrium Fisher's exact test *P* value of lower than  $1 \times 10^{-6}$ , or call rate lower than 99% were removed. The imputation was done using IMPUTE v2.0 (Howie et al., 2009) with reference haplotypes of HapMap 2 release 22 (<http://hapmap.ncbi.nlm.nih.gov/downloads/phasing/2007-08%20rel22/phased/>).

### Lipid Measurements

In all twin cohorts, we measured HDL, LDL, TC, and TG according to standard enzymatic methods. For a subset of twin-pairs, we also measured apolipoproteins A-I, A-II, B, and E, using immunoassays as described in Beekman et al. (2002).

### Statistical Analysis

Because of a skewed distribution, TG and apolipoprotein A1, A2, B, and E values were log-transformed and, for all of phenotypes, outliers deviating more than three times the standard deviation from the trait mean were removed. Then means and absolute differences in lipid and apolipoprotein values were calculated for each pair. Both intra-pair differences and means were transformed into inverse normal distribution by rank for further analysis. Trait values were

then adjusted for age (and differences additionally by mean effect) using a linear regression model. Association analyses between the imputed SNP probabilities and adjusted differences and means were conducted using SNPtest (Marchini et al., 2007). Cohort-specific results were then combined using the inverse variance meta-analysis method function in MetABEL (<http://mga.bionet.nsc.ru/~yurii/ABEL/>) package for statistical analysis program, R (<http://cran.r-project.org/>). We applied minor allele frequency threshold of 5% and imputation quality threshold of proper info >0.4 before meta-analysis.

### Replication Samples

The Finnish replication twin set is a part of the Nicotine Addiction Genetics – Finland (FTC/NAG-FIN) study. FTC/NAG-FIN sample was ascertained from the Finnish Twin Cohort study consisting of adult twins born between 1938 and 1957 ([www.twinstudy.helsinki.fi](http://www.twinstudy.helsinki.fi)). Based on earlier health questionnaires, the twin-pairs concordant for ever-smoking were identified and recruited along with their family members (mainly siblings) for the NAG-FIN study ( $N = 2,265$ ) as part of the consortium, including Finland, Australia, and United States (Broms et al., 2012; Loukola et al., 2008). Data collection took place between 2001 and 2005. The dataset was genotyped with Illumina human 670K chip in Wellcome Trust Sanger Institute.

The additional Swedish twins were from the TwinGene project, which is a part of the Swedish Twin registry. Twins born before 1958 were contacted to participate in a simple health check-up, with measurement of height, weight, waist and hip circumference, and blood pressure. Health and medication data were collected from self-reported questionnaires, and blood sampling materials were mailed to the subjects who then went to a local healthcare center for blood sampling for subsequent DNA extraction, serum collection, and clinical chemistry tests. Dataset was genotyped at the SNP Technology Platform in Uppsala University with Illumina OmniExpress 700K chip.

## Results

The cohort-specific characteristics for the MZ twin sets are presented in Supplementary Table S1 online. The average age varies among the twin cohorts, being lowest in the Netherlands (mean age 33.6 years) and highest in Sweden (mean age 71.8 years). The MZ within-pair correlations in our dataset for lipids and lipoproteins are presented in Table 1. The within-pair correlations vary between the cohorts by their mean age as can be seen, for example, for the total cholesterol measures: For the oldest cohorts, Finnish and Swedish, the intra-pair correlations are 0.3 and 0.4, while for the younger cohorts the correlation is around 0.6–0.7 (Table 1). For this, as well as for the known age-trend in lipid levels, all the statistical analyses presented below are controlled for age.

**TABLE 1**  
Within-Pair Correlations

Cohort	HDL	LDL	TC	TG	APOA1	APOA2	APOB	APOE
Australia	0.611	0.587	0.597	0.515	0.590	0.462	0.690	0.570
UK	0.670	0.688	0.687	0.657	na	na	na	na
Sweden	0.634	0.355	0.306	0.471	0.638	na	0.418	na
Finland	0.756	0.495	0.443	0.519	0.440	0.263	0.409	na
Denmark	0.593	0.768	0.769	0.558	0.777	na	0.649	0.314
Netherlands	0.629	0.664	0.669	0.537	0.847	0.528	0.801	0.942
Combined mean	0.651	0.595	0.589	0.552	0.638	0.461	0.598	0.531

Note: The within MZ twin-pair correlations for high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, total cholesterol (TC), logarithmized triglyceride (TG), logarithmized apolipoproteins A1 (APOA1), A2 (APOA2), B (APOB), and E (APOE).  
No adjustment for age was applied.  
The combined means have been calculated using weights based on the size of the cohort.  
na: not available.

**TABLE 2**  
Intra-Pair Difference Association Meta-Analysis Results

Phenotype	SNP	Chromosome	Position Build 36	Effect allele	Minor allele frequency	Beta (SE)	N	P value	Gene
HDL	rs2483058	1	204693451	C	0.41	0.203 (0.037)	1,634	$3.98 \times 10^{-8}$	SRGAP2 (intronic)
HDL	rs9242	1	204704018	C	0.47	0.191 (0.036)	1,634	$1.08 \times 10^{-7}$	SRGAP2 (untranslated)
TC	rs17826288	3	109178193	G	0.38	-0.178 (0.036)	1,720	$8.16 \times 10^{-7}$	66 kb downstream of CD47
APOA2	rs6032680	20	44215538	G	0.43	-0.394 (0.079)	696	$5.82 \times 10^{-7}$	20 kb downstream of CDH22

Note: SNPs having  $P$  value  $< 1 \times 10^{-6}$  in the association analysis of inverse normal transformed absolute intra pair differences.  
SE: standard error of effect estimate, beta; N: number of individuals; HDL: high-density lipoprotein; TC: total cholesterol; APOA2: apolipoprotein A2.

**TABLE 3**  
Effects of the Genome-Wide Significant SNPs in SRGAP2 Gene in the Six Twin Cohorts

SNP	Australia Beta (SE)	UK Beta (SE)	Sweden Beta (SE)	Finland Beta (SE)	Denmark Beta (SE)	The Netherlands Beta(SE)
rs2483058	0.140 (0.073)	0.213 (0.078)	0.227 (0.084)	0.219 (0.133)	0.265 (0.119)	0.216 (0.094)
rs9242	0.122 (0.070)	0.226 (0.076)	0.246(0.085)	0.164 (0.135)	0.234 (0.115)	0.180 (0.089)

Note: Cohort wise results of the SNPs having  $P$  value  $< 5 \times 10^{-7}$  in the association analysis of inverse normal transformed absolute intra pair differences. SE: standard error of the effect estimate, beta.

We used the imputed genome-wide scan to identify genes associated with variability of lipids and apolipoproteins (see Figure 1). One locus having two SNPs with genome-wide significant effect ( $P < 5 \times 10^{-8}$ ) on HDL variability was found (Table 2, Figure 2) and the effects were consistent across our six twin samples and in a joint analysis (Table 3). The variant with stronger association is located in an intron of the *SRGAP2* gene on chromosome 1.

We tried to replicate this association in two additional MZ twin datasets that showed positive effect for the C allele as in the initial analysis (Table 4), but with 50% smaller effect size. When combining results from the initial and replication analyses, the  $P$  value did not reach genome-wide significance (combined  $P$  value =  $8.7 \times 10^{-7}$ ). In addition to twin datasets, we followed the finding in a population cohort of 15,874 non-twin individuals with Bartlett's test of homogeneity of variance, but could not find support for the finding (combined  $P = 0.56$ , data not shown).

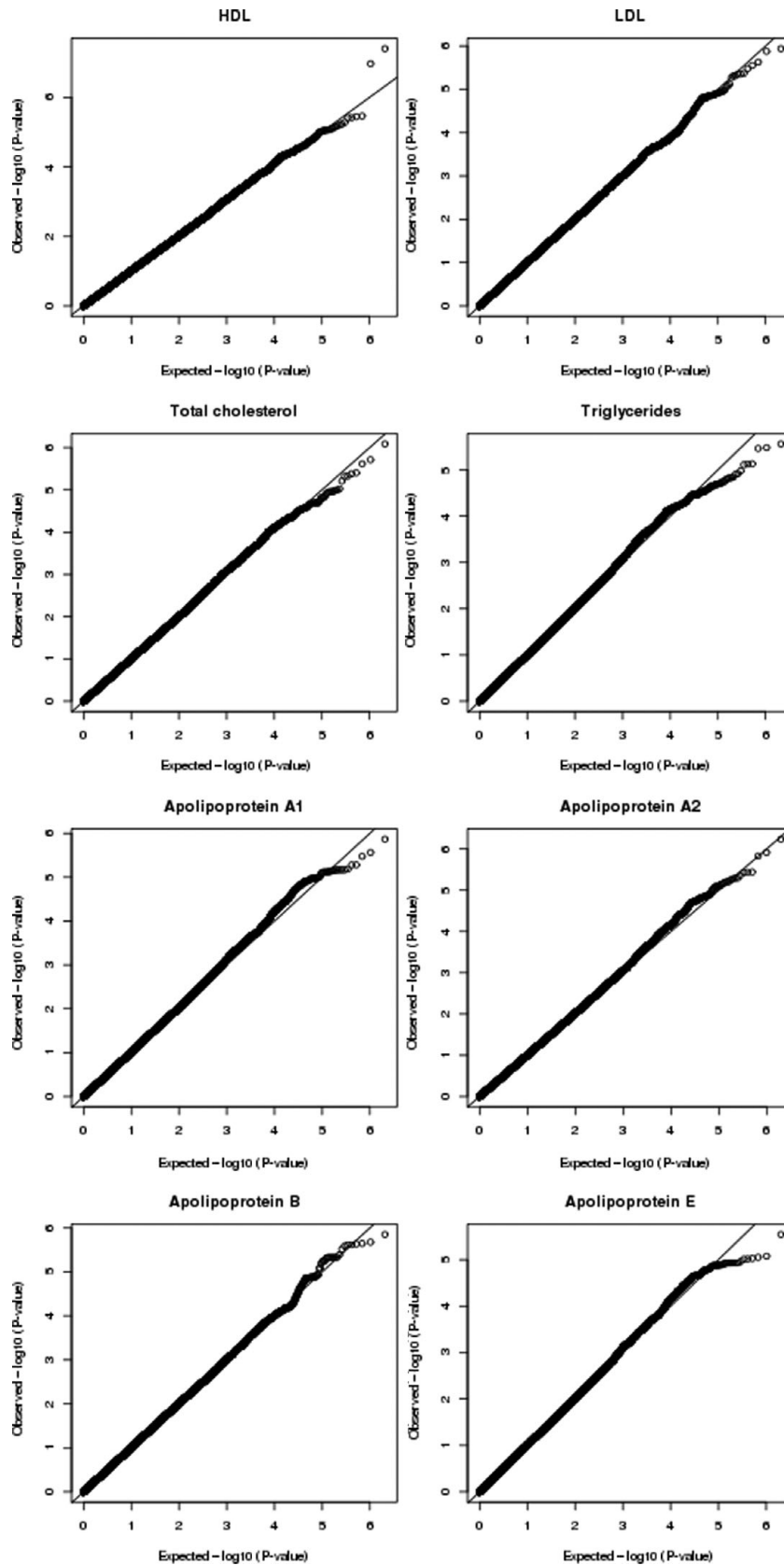
In addition to the one genome-wide significant hit, there was one suggestive hit ( $P = 8.16 \times 10^{-7}$ ) for TC variability,

*rs17826288*, which is positioned 66 kb downstream of the *CD47* gene. Follow up of this locus in the population-based cohort resulted in a combined Bartlett's test  $P$ -value of 0.005, but the variances increased with the G allele, while in the MZGWA sample the variability decreases.

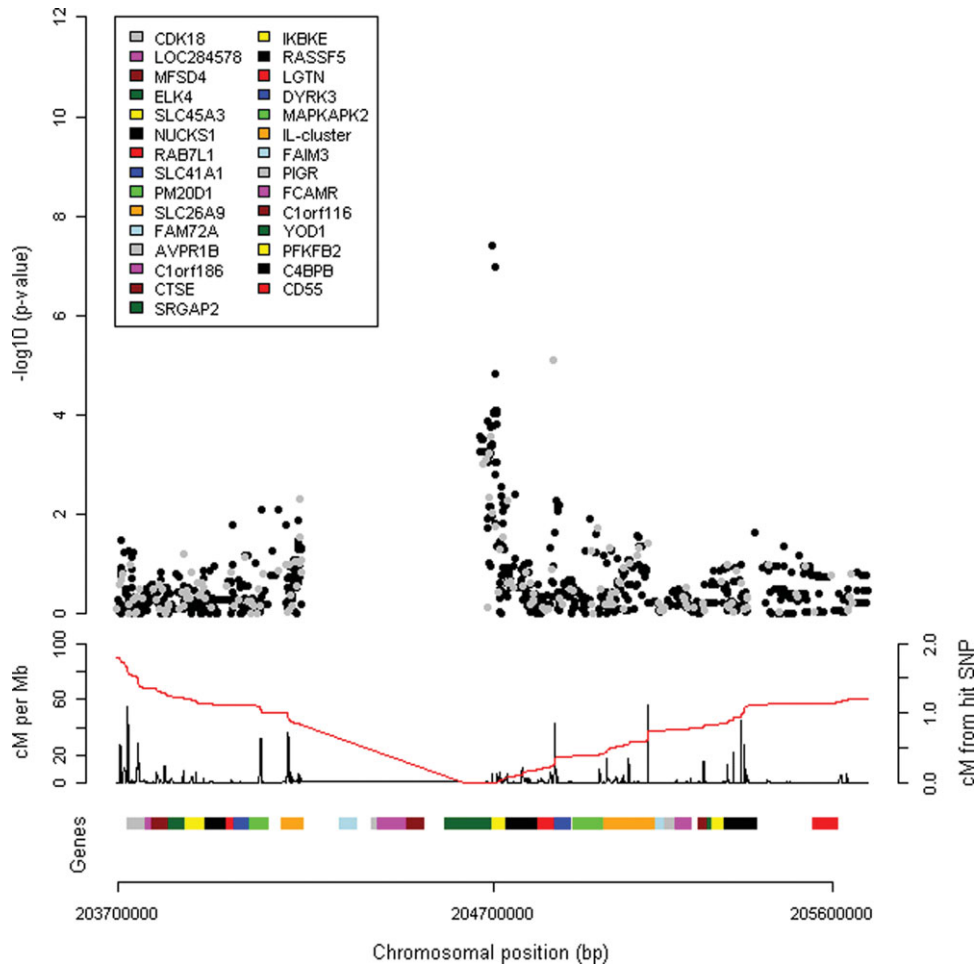
In addition, the aim of this study was to identify possible genes associated with apolipoprotein differences. No genome-wide significant hits were found, but we did find one suggestive hit associated with APOA2 intra-pair difference with a  $P$  value of  $5.82 \times 10^{-7}$ . This SNP, *rs6032680*, is 20 kb downstream of *CDH22* gene (Table 2). We could not replicate this finding, as the additional twin datasets and population-based cohort did not have measurement of apolipoproteins.

Results from the apolipoprotein level association screening are presented in Table 5. For the *APOE/APOC1* locus, there was one SNP that had a genome-wide significant association for APOB and APOE. In addition, we found one new genome-wide significant locus for apolipoprotein levels, *DOCK1*, which had an association with APOA2 levels. There was also suggestive deviation from the expected  $P$





**FIGURE 1**  
 Quantile–quantile plot of all eight variability association analyses. Deviation from expected could be seen in HDL–cholesterol only.  
 QQ-PLOTS OF VARIABILITY ASSOCIATION ANALYSES



**FIGURE 2**

(Colour online) The upper panel shows  $-\log_{10}(p\text{-value})$  where the black dots are imputed and gray dots genotyped SNPs. In the second panel, the black line is the recombination rate (cM/Mb) and the red line is the distance from the best SNP measured as cM. The recombination rate and genetic positions are from [http://hapmap.ncbi.nlm.nih.gov/downloads/recombination/2008-03\\_rel22\\_B36/rates/](http://hapmap.ncbi.nlm.nih.gov/downloads/recombination/2008-03_rel22_B36/rates/). REGIONAL ASSOCIATION PLOT OF THE *SRGAP2* LOCUS

**TABLE 4**

Replication of *rs2483058* Association on HDL-Variability in Additional Twin Datasets

Cohort	Sex	Effect allele	Beta	SE	N	P value
Finnish Twin cohort	Male	C	-0.012	0.171	65	0.944
Finnish Twin cohort	Female	C	0.140	0.156	87	0.369
Swedish Twin Registry	Male	C	0.078	0.047	562	0.100
Swedish Twin Registry	Female	C	0.032	0.046	547	0.483
Meta		C	0.098	0.032	1261	0.0020

Note: HDL-variability association results for *rs2483058* in two additional monozygotic twin datasets. SE: standard error of effect estimate, beta. N: number of individuals.

**TABLE 5**

Genome-Wide Significant Association Analysis Results on Apolipoprotein Levels

Locus	SNP	Trait	Chromosome	Position of the SNP Build 36	Effect allele	N	Beta (SE)	P value
<i>DOCK1</i>	<i>rs9733352</i>	APOA2	10	128667380	C	444	0.060 (0.011)	$4.03 \times 10^{-8}$
<i>APOE/APOC1</i>	<i>rs445925</i>	APOB	19	50107480	G	875	0.149 (0.019)	$1.65 \times 10^{-14}$
		APOE				573	-0.229 (0.035)	$4.18 \times 10^{-11}$

Note: SNPs having P value of lower than  $5 \times 10^{-8}$  in the association analysis of intra pair means of apolipoproteins. N: number of individuals; SE: standard error of the effect estimate, beta; APOA2: apolipoprotein A2; APOB: apolipoprotein B; APOE = apolipoprotein E.

values (supplementary Figure S1 online) for APOA1, but no genome-wide significant loci were found.

## Discussion

It has been shown previously (Visscher & Posthuma, 2010) that the MZ difference approach has more power than the population variance method when the within-pair correlation of twins is higher than 0.3, which is the case for both circulating blood lipids and apolipoproteins (Table 1). Following the formulae in the study by Visscher and Posthuma (2010), with a within-pair correlation of 0.651 (0.605 for CC homozygous pairs), which was derived from the MZ dataset, we get a multiplicative effect of 1.17 in population variances. For a hit with a multiplicative effect of 1.2 in variances, the relative efficiency is approximately two to three times higher for the twin approach and the number of MZ pairs used in this (1,720) study corresponds to a population sample of 11,000.

Large international efforts combining results from multiple cohorts can achieve the power to detect variability of  $G \times E$  associations by increasing the number of individuals in the analysis, but at the same time, more variability is induced by population-specific effects. One of the challenges in gene–environment interaction mapping with quantitative traits is the amount of extra variability in the analysis because of the measurement error of the environmental factor. However, when mapping variability genes instead of  $G \times E$  interactions, this is not an issue because no specific environmental factor is assumed.

We can infer from twin data that the differential variability is caused by differential sensitivity to the environment or lifestyle and not by other interacting genes, since the MZ twins are essentially genetically identical at the sequence level. With our genome-wide MZ twin dataset we showed initial evidence for a locus associated with twins being different in their HDL values depending on their genotypes in *SRGAP2* gene. The *SRGAP2* has been connected to multiple different cancers, but there is no known connection between the gene and lipid metabolism.

Because of the restricted availability of MZ twin datasets, we additionally tried to replicate the genome-wide significant finding in population based cohorts but could not find support. This failed replication could be due to false positive finding or power issues in homogeneity of variance analysis with Bartlett's test. The suggestive TC variability locus, *CD47*, showed some evidence of variability association in the population cohorts, but in opposite direction.

By the analysis of MZ-twin apolipoprotein levels, we were able to replicate one previously published lipid gene, but also found a new locus associated with APOA2. As this is one of the first GWA analyses studying apolipoprotein fractions in addition to lipids, further replication studies for the new locus, *DOCK1*, are called for.

The mechanisms of joint actions of the gene and lifestyle factors are unknown, but these suggestive results provide

basis for further functional and epidemiological studies. Further, detailed consideration of environmental factors affecting lipid levels as well as developmental changes over the lifespan need to be incorporated in the analysis models.

## Author Contributions

MP, A-CS, AP, JK, KOK, NLP, DIB, TS, NGM, SR, and LP conceived and designed the experiments. JBW, MP, PMV, GWM, MF, GW, EJCdG, PKEM, KC, TIAS, KHP, TR, KS, EW, UJ, A-CS, and AP performed the experiments. IS, MP, JM, IR, and SR analyzed the data. SR, MP, IS, NGM, PMV, GWM, MF, PKEM, KS, EW, UL, A-CS, AP, DIB, and LP contributed reagents/materials/analysis tools. All authors wrote and approved the final version of the paper.

Supplementary Material is available on the Cambridge Journals Online website.

## Acknowledgments

The GenomEUtwin project is supported by the European Commission under the program “Quality of Life and Management of the Living Resources” of 5th Framework Programme (QLG2-CT-2002–01254).

The Swedish Twin Registry was supported by the US National Institutes of Health (AG028555, AG08724, AG04563, AG10175, AG08861), the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research, the Royal Swedish Academy of Science, and ENGAGE (within the European Union Seventh Framework Programme, HEALTH – F4-2007-201413).

The Danish Twin Registry has been supported by the Danish Medical Research Council, the Danish Diabetes Foundation, the Danish Heart Association, and the Novo Nordic Foundation.

The TWINSUK study was funded by the Wellcome Trust (Grant ref. 079771); European Community's Framework 6 Project EUroClot. The study also received support from the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. Timothy D Spector is an NIHR senior investigator and is a holder of an ERC Advanced Principal Investigator award.

The Australian Twin Cohort was funded by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), and the US National Institutes of Health (AA07535, AA10248, AA11998, AA13320, AA13321, AA13326, AA14041, AA17688, DA12854, MH66206).

The Netherlands Twin Registry (NTR) has been supported by the Netherlands Scientific Organization (911–09–032, 480–04–004, CMSB NOW Genomics, NWO–SPI 56–464–1419, NBIC/BioAssist/RK/2008.024), BBMRI–6 NL,



Institute for Health and Care Research (EMGO+), the European Union (EU/WLRT–2001–01254, ERC Advanced 230374).

The Finnish Twin Cohort Study was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grant numbers 213506, 129680), the Finnish Foundation for Cardiovascular Research, the Finnish Academy (grant numbers 129494, 139635, and 251217), and the Sigrid Juselius Foundation.

Platform in Uppsala is funded by Uppsala University and Uppsala University Hospital and the Swedish Wallenberg Foundation.

SR was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (213506 and 129680), Academy of Finland (251217), the Finnish foundation for Cardiovascular Research, and the Sigrid Juselius Foundation. MP is partly financially supported for this work by the Finnish Academy SALVE program “Pubgensense” 129322 and by grants from Finnish Foundation for Cardiovascular Research. KP was funded by Helsinki University Central Hospital grants. TR has been funded from the Academy of Finland Research Programme on Ageing, Finnish Ministry of Education.

We thank the staff of the SNP Platform for their contributions to genotyping.

## References

- Aulchenko, Y. S., Ripatti, S., Lindqvist, I., Boomsma, D., Heid, I. M., Pramstaller, P. P., Penninx, B. W., Janssens, A. C., Wilson, J. F., Spector, T., Martin, N. G., Pedersen, N. L., Kyvik, K. O., Kaprio, J., Hofman, A., Freimer, N. B., Jarvelin, M. R., Gyllensten, U., Campbell, H., Rudan, I., Johansson, A., Marroni, F., Hayward, C., Vitart, V., Jonasson, I., Pattaro, C., Wright, A., Hastie, N., Pichler, I., Hicks, A. A., Falchi, M., Willemsen, G., Hottenga, J. J., de Geus, E. J., Montgomery, G. W., Whitfield, J., Magnusson, P., Saharinen, J., Perola, M., Silander, K., Isaacs, A., Sijbrands, E. J., Uitterlinden, A. G., Witteman, J. C., Oostra, B. A., Elliott, P., Ruokonen, A., Sabatti, C., Gieger, C., Meitinger, T., Kronenberg, F., Döring, A., Wichmann, H. E., Smit, J. H., McCarthy, M. I., van Duijn, C. M., Peltonen, L., & ENGAGE Consortium. (2008). Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature Genetics*, *41*, 47–55.
- Beekman, M., Heijmans, B. T., Martin, N. G., Pedersen, N. L., Whitfield, J. B., DeFaire, U., van Baal, G. C., Snieder, H., Vogler, G. P., Slagboom, P. E., & Boomsma, D. I. (2002). Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Research*, *5*, 87–97.
- Broms, U., Wedenoja, J., Largeau, M. R., Korhonen, T., Pitkaniemi, J., Kesitalo-Vuokko, K., Häppölä, A., Heikkilä, K. H., Heikkilä, K., Ripatti, S., Sarin, A. P., Salminen, A. P., Paunio, T., Pergadia, M. L., Madden, P. A., Kaprio, J., & Loukola, A. (2012). Analysis of detailed phenotype profiles reveals CHRNA5-CHRNA3-CHRNA4 gene cluster association with several nicotine dependence traits. *Nicotine & Tobacco Research*, *14*, 720–733.
- Friedlander, Y., Austin, M. A., Newman, B., Edwards, K., Mayer-Davis, E. I., & King, M. C. (1997). Heritability of longitudinal changes in coronary-heart-disease risk factors in women twins. *American Journal of Human Genetics*, *60*, 1502–1512.
- Hannelius, U., Gherman, L., Mäkelä, V. V., Lindstedt, A., Zuchelli, M., Lagerberg, C., Tybring, G., Kere, J., & Lindgren, C. M. (2007). Large-scale zygosity testing using single nucleotide polymorphisms. *Twin Research and Human Genetics*, *10*, 604–625.
- 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 Genetics*, *5*, e1000529.
- Kannel, W. B., Kagan, A., Revotskie, N., & Stokes, J. I. (1961). Factors of risk in the development of coronary heart disease—six year follow-up experience. The Framingham Study. *Annals of Internal Medicine*, *55*, 33–50.
- Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burt, N. P., Rieder, M. J., Cooper, G. M., Roos, C., Voight, B. F., Havulinna, A. S., Wahlstrand, B., Hedner, T., Corella, D., Tai, E. S., Ordovas, J. M., Berglund, G., Vartiainen, E., Jousilahti, P., Hedblad, B., Taskinen, M. R., Newton-Cheh, C., Salomaa, V., Peltonen, L., Groop, L., Altshuler, D. M., & Orholm-Melander, M. (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genetics*, *40*, 189–197.
- Kooner, J. S., Chambers, J. C., Aguilar-Salinas, C. A., Hinds, D. A., Hyde, C. L., Warnes, G. R., Gomez Perez, F. J., Frazer, K. A., Elliott, P., Scott, J., Milos, P. M., Cox, D. R., & Thompson, J. F. (2008). Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nature Genetics*, *40*, 149–151.
- Loukola, A., Broms, U., Maunu, H., Widén, E., Heikkilä, K., Siivola, M., Salo, A., Pergadia, M. L., Nyman, E., Sarmalisto, S., Perola, M., Agrawal, A., Heath, A. C., Martin, N. G., Madden, P. A., Peltonen, L., & Kaprio, J. (2008). Linkage of nicotine dependence and smoking behavior on 10q, 7q and 11p in twins with homogeneous genetic background. *Pharmacogenomics Journal*, *8*, 209–219.
- Magnus, P., Berg, K., Borreson, A. L., & Nance, W. E. (1981). Apparent influence of marker genotypes on variation in serum cholesterol in monozygotic twins. *Clinical Genetics*, *19*, 67–70.
- Marchini, J., Howie, B. N., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genetics*, *39*, 906–913.
- Miller, N. E., & Miller, M. G. (1975). Letter to the Editor: High-density lipoprotein and atherosclerosis. *Lancet*, *305*, 1033.
- Nyholt, D. R. (2006). On the probability of dizygotic twins being concordant for two alleles at multiple polymorphic loci. *Twin Research and Human Genetics*, *9*, 194–197.
- Peltonen, L., & GenomEUtwin project. (2003). A Strategy to Identify Genetic Influences on Health and Disease. *Twin Research*, *6*, 354–360.

- Pilia, G., Chen, W. M., Scuteri, A., Orru, M., Albai, G., Dei, M., Lai, S., Usala, G., Lai, M., Loi, P., Mamei, C., Vacca, L., Deiana, M., Olla, N., Masala, M., Cao, A., Najjar, S. S., Terracciano, A., Nedorezov, T., Sharov, A., Zonderman, A. B., Abecasis, G. R., Costa, P., Lakatta, E., & Schlessinger, D. (2006). Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genetics*, 2, e132.
- Stemmers, F. J., Chang, W., Lee, G., Barker, D. L., Shen, R., & Gunderson, K. L. (2006). Whole-genome genotyping with the single-base extension assay. *Nature Methods*, 3, 31–33.
- Surakka, I., Isaacs, A., Karssen, L., Laurila, P.-P., Middelberg, R., Tikkanen, E., Ried, J., Lamina, C., Mangino, M., Igl, W., Hottenga, J., Lagou, V., van der Harst, P., Mateo Leach, I., Esko, T., Kutalik, Z., Wainwright, N., Struchalin, M., Sarin, A.-P., Kangas, A., Viikari, J., Perola, M., Rantanen, T., Petersen, A.-K., Soyninen, P., Johansson, Å., Soranzo, N., Heath, A., Papamarkou, T., Prokopenko, I., Tönjes, A., Kronenberg, F., Döring, A., Rivadeneira, F., Montgomery, G., Whitfield, J., Kähönen, M., Lehtimäki, T., Freimer, N., Willemsen, G., de Geus, E., Palotie, A., sandhu, M., Waterworth, D., Metspalu, A., Stumvoll, M., Uitterlinden, A., Jula, A., Navis, G., Wijmenga, C., Wolffenbuttel, B., Taskinen, M., Ala-Korpela, M., Kaprio, J., Kyvik, K., Boomsma, D., Pedersen, N., Gyllensten, U., Wilson, J., Rudan, I., Campbell, H., Pramstaller, P., Spector, T., Witteman, J., Eriksson, J., Salomaa, V., Oostra, B., Raitakari, O., Wichmann, H., Gieger, C., Jarvelin, M.-R., Martin, N., Hofman, A., McCarthy, M., Peltonen, L., Van Duijn, C., Aulchenko, Y., Ripatti, S., & ENGAGE Consortium (2011). A genome-wide screen for interactions reveals a new locus on 4p15 modifying the effect of waist-to-hip ratio on total cholesterol. *PLoS Genetics*, 7, e1002333.
- Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P., Ripatti, S., Chasman, D. I., Willer, C. J., Johansen, C. T., Fouchier, S. W., Isaacs, A., Peloso, G. M., Barbalic, M., Ricketts, S. L., Bis, J. C., Aulchenko, Y. S., Thorleifsson, G., Feitosa, M. F., Chambers, J., Orho-Melander, M., Melander, O., Johnson, T., Li, X., Guo, X., Li, M., Shin Cho, Y., Jin Go, M., Jin Kim, Y., Lee, J. Y., Park, T., Kim, K., Sim, X., Tzee-Hee Ong, R., Croteau-Chonka, D. C., Lange, L. A., Smith, J. D., Song, K., Hua Zhao, J., Yuan, X., Luan, J., Lamina, C., Ziegler, A., Zhang, W., Zee, R. Y., Wright, A. F., Witteman, J. C., Wilson, J. F., Willemsen, G., Wichmann, H. E., Whitfield, J. B., Waterworth, D. M., Wareham, N. J., Waeber, G., Vollenweider, P., Voight, B. F., Vitart, V., Uitterlinden, A. G., Uda, M., Tuomilehto, J., Thompson, J. R., Tanaka, T., Surakka, I., Stringham, H. M., Spector, T. D., Soranzo, N., Smit, J. H., Sinisalo, J., Silander, K., Sijbrands, E. J., Scuteri, A., Scott, J., Schlessinger, D., Sanna, S., Salomaa, V., Saharinen, J., Sabatti, C., Ruukonen, A., Rudan, I., Rose, L. M., Roberts, R., Rieder, M., Psaty, B. M., Pramstaller, P. P., Pichler, I., Perola, M., Penninx, B. W., Pedersen, N. L., Pattaro, C., Parker, A. N., Pare, G., Oostra, B. A., O'Donnell, C. J., Nieminen, M. S., Nickerson, D. A., Montgomery, G. W., Meitinger, T., McPherson, R., McCarthy, M. I., McArdle, W., Masson, D., Martin, N. G., Marroni, F., Mangino, M., Magnusson, P. K., Lucas, G., Luben, R., Loos, R. J., Lokki, M. L., Lettre, G., Langenberg, C., Launer, L. J., Lakatta, E. G., Laaksonen, R., Kyvik, K. O., Kronenberg, F., König, I. R., Khaw, K. T., Kaprio, J., Kaplan, L. M., Johansson, A., Jarvelin, M. R., Janssens, A. C., Ingelsson, E., Igl, W., Kees Hovingh, G., Hottenga, J. J., Hofman, A., Hicks, A. A., Hengstenberg, C., Heid, I. M., Hayward, C., Havulinna, A. S., Hastie, N. D., Harris, T. B., Haritunians, T., Hall, A. S., Gyllensten, U., Guiducci, C., Groop, L. C., Gonzalez, E., Gieger, C., Freimer, N. B., Ferrucci, L., Erdmann, J., Elliott, P., Ejebe, K. G., Döring, A., Dominiczak, A. F., Demissie, S., Deloukas, P., de Geus, E. J., de Faire, U., Crawford, G., Collins, F. S., Chen, Y. D., Caulfield, M. J., Campbell, H., Burt, N. P., Bonnycastle, L. L., Boomsma, D. I., Boekholdt, S. M., Bergman, R. N., Barroso, I., Bandinelli, S., Ballantyne, C. M., Assimes, T. L., Quertermous, T., Alshuler, D., Seielstad, M., Wong, T. Y., Tai, E. S., Feranil, A. B., Kuzawa, C. W., Adair, L. S., Taylor, H. A. Jr, Borecki, I. B., Gabriel, S. B., Wilson, J. G., Holm, H., Thorsteinsdottir, U., Gudnason, V., Krauss, R. M., Mohlke, K. L., Ordovas, J. M., Munroe, P. B., Kooner, J. S., Tall, A. R., Hegele, R. A., Kastelein, J. J., Schadt, E. E., Rotter, J. I., Boerwinkle, E., Strachan, D. P., Mooser, V., Stefansson, K., Reilly, M. P., Samani, N. J., Schunkert, H., Cupples, L. A., Sandhu, M. S., Ridker, P. M., Rader, D. J., van Duijn, C. M., Peltonen, L., Abecasis, G. R., Boehnke, M., & Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, 466, 707–713.
- Thorn, J. A., Needham, E. W., Mattu, R. K., Stocks, J., & Galton, D. J. (1998). The Ser447-Ter mutation of the lipoprotein lipase gene relates to variability of serum lipid and lipoprotein levels in monozygotic twins. *Journal of Lipid Research*, 39, 437–441.
- Visscher, P. M., & Duffy, D. L. (2006). The value of relatives with phenotypes but missing genotypes in association studies for quantitative traits. *Genetic Epidemiology*, 30, 30–36.
- Visscher, P., & Posthuma, D. (2010). Statistical power to detect genetic loci affecting environmental sensitivity. *Behavioural Genetics*, 40, 728–733.
- Willer, C. J., Sanna, S., Jackson, A. U., Scuteri, A., Bonnycastle, L. L., Clarke, R., Heath, S. C., Timpson, N. J., Najjar, S. S., Stringham, H. M., Strait, J., Duren, W. L., Maschio, A., Busonero, F., Mulas, A., Albai, G., Swift, A. J., Morken, M. A., Narisu, N., Bennett, D., Parish, S., Shen, H., Galan, P., Meneton, P., Hercberg, S., Zelenika, D., Chen, W. M., Li, Y., Scott, L. J., Scheet, P. A., Sundvall, J., Watanabe, R. M., Nagaraja, R., Ebrahim, S., Lawlor, D. A., Ben-Shlomo, Y., Davey-Smith, G., Shuldiner, A. R., Collins, R., Bergman, R. N., Uda, M., Tuomilehto, J., Cao, A., Collins, F. S., Lakatta, E., Lathrop, G. M., Boehnke, M., Schlessinger, D., Mohlke, K. L., & Abecasis, G. R. (2008). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nature Genetics*, 40, 161–169.