Human Molecular Genetics, 2012, Vol. 21, No. 6 **1433–1443**
doi:10.1093/hmg/ddr580
Advance Access published on December 8, 2011

Genetic associations with lipoprotein subfractions provide information on their biological nature

Ann-Kristin Petersen[1], Klaus Stark[7], Muntaser D. Musameh[8,9], Christopher P. Nelson[8,9], Werner Römisch-Margl[2], Werner Kremer[10,11], Johannes Raffler[2,12], Susanne Krug[13], Thomas Skurk[13], Manuela J. Rist[14], Hannelore Daniel[14], Hans Hauner[13], Jerzy Adamski[3,15], Maciej Tomaszewski[8,9], Angela Döring[4,5], Annette Peters[5], H.-Erich Wichmann[4,16,17], Bernhard M. Kaess[7,8], Hans Robert Kalbitzer[10,11], Fritz Huber[10], Volker Pfahlert[10], Nilesh J. Samani[8,9], Florian Kronenberg[18], Hans Dieplinger[18], Thomas Illig[6,20], Christian Hengstenberg[7], Karsten Suhre[2,12,19], Christian Gieger[1,*] and Gabi Kastenmüller[2]

[1]Institute of Genetic Epidemiology, [2]Institute of Bioinformatics and Systems Biology, [3]Institute of Experimental Genetics, Genome Analysis Center, [4]Institute of Epidemiology I, [5]Institute of Epidemiology II and [6]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg 85764, Germany, [7]Klinik und Poliklinik für Innere Medizin II, University of Regensburg, Regensburg 93053, Germany, [8]Department of Cardiovascular Sciences, University of Leicester, Leicester LE3 9QP, UK, [9]Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester LE3 9QP, UK, [10]LipoFIT Analytic GmbH, Regensburg 93053, Germany, [11]Institute of Biophysics and Physical Biochemistry and Centre of Magnetic Resonance in Chemistry and Biomedicine, University of Regensburg, Regensburg 93040, Germany, [12]Faculty of Biology, Ludwig-Maximilians-Universität, Planegg-Martinsried 82152, Germany, [13]Else Kröner-Fresenius-Centre for Nutritional Medicine, [14]Molecular Nutrition Unit and [15]Lehrstuhl für Experimentelle Genetik, Technische Universität München, Freising-Weihenstephan 85350, Germany, [16]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, München 81377, Germany, [17]Grosshadern, Klinikum, München 81377, Germany, [18]Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck 6020, Austria, [19]Department of Physiology and Biophysics, Weill Cornell Medical College in Qatar, Education City-Qatar Foundation, Doha, Qatar and [20]Hannover Unified Biobank, Hannover Medical School, Hannover 30625, Germany

Received July 11, 2011; Revised November 10, 2011; Accepted December 5, 2011

Adverse levels of lipoproteins are highly heritable and constitute risk factors for cardiovascular outcomes. Hitherto, genome-wide association studies revealed 95 lipid-associated loci. However, due to the small effect sizes of these associations large sample numbers (>100 000 samples) were needed. Here we show that analyzing more refined lipid phenotypes, namely lipoprotein subfractions, can increase the number of significantly associated loci compared with bulk high-density lipoprotein and low-density lipoprotein analysis in a study with identical sample numbers. Moreover, lipoprotein subfractions provide novel insight into the human lipid metabolism. We measured 15 lipoprotein subfractions (L1–L15) in 1791 samples using ^{1}H-NMR (nuclear magnetic resonance) spectroscopy. Using cluster analyses, we quantified inter-relationships among lipoprotein subfractions. Additionally, we analyzed associations with subfractions at known lipid loci. We identified five distinct groups of subfractions: one (L1) was only marginally captured by serum lipids and therefore extends our knowledge of lipoprotein biochemistry. During a lipid-tolerance test, L1 lost its special position. In the association analysis, we found that eight loci (*LIPC*, *CETP*, *PLTP*, *FADS1-2-3*,

*To whom correspondence should be addressed at: Institute of Genetic Epidemiology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany. Tel: +49 8931874106; Fax: +49 8931873380; Email: christian.gieger@helmholtz-muenchen.de

***SORT1*, *GCKR*, *APOB*, *APOA1*) were associated with the subfractions, whereas only four loci (*CETP*, *SORT1*, *GCKR*, *APOA1*) were associated with serum lipids. For *LIPC*, we observed a 10-fold increase in the variance explained by our regression models. In conclusion, NMR-based fine mapping of lipoprotein subfractions provides novel information on their biological nature and strengthens the associations with genetic loci. Future clinical studies are now needed to investigate their biomedical relevance.**

INTRODUCTION

Genome-wide association studies (GWAS) are currently the state-of-the-art method to reveal new genetic risk loci for quantitative traits. At present, 95 associated common variants are reported for high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), triglycerides (TG) and total cholesterol (TC) (1). These loci explain 10–12% of the total variance of serum lipids. Extreme levels of serum lipids are a major risk factor for cardiovascular outcomes such as coronary artery disease, myocardial infarction and stroke (2,3). While the contribution of LDL-C to the development of coronary artery disease is well documented, the role of other lipoprotein fractions (including HDL-C) in atherosclerosis and its clinical manifestations is less well understood (4–6). For example, the torcetrapib failure revealed the complexity of the HDL metabolism and implicated that further research on HDL and HDL fractions is needed (7). To obtain a more detailed view of the lipid metabolism, subfractions of lipoproteins can be analyzed. Among others, conventional ^{1}H-NMR (nuclear magnetic resonance) spectroscopy was used by two research groups (8,9) to develop a high-throughput lipoprotein quantification method. Recent studies (10) have led to a modified software using a 400 MHz NMR lipoprotein analyzer that measures nine lipoprotein subclasses with high reproducibility (11–13). Using this technology, Chasman *et al.* (11) conducted a GWAS with the aim of finding new lipoprotein loci to better understand genetic associations with the lipoprotein metabolism. The accuracy of conventional ^{1}H-NMR spectroscopy increases with the magnetic field strength. Therefore, Ala-Korpela *et al.* (14) recorded conventional ^{1}H-NMR spectra at a magnetic field strength of 600 MHz and were able to quantify 11 lipoprotein subclasses with high reproducibility (15). The technology used in our study permits one to quantify a set of 15 lipoprotein subclass coefficients (L1–L15) by diffusion-weighted 600 MHz ^{1}H-NMR spectroscopy, which allows the measurement of diffusion in complex mixtures. The profiles derived from this technology have been investigated in a number of studies (16–21).

We analyzed the 15 lipoprotein subfractions obtained from diffusion-weighted NMR measurements. These NMR-derived lipoprotein subfractions were separated by their apparent differences in the methyl and methylene signals in diffusion-weighted NMR experiments. Usually, the lipoprotein subfractions obtained by NMR are ordered according to their diffusion coefficients, which in turn are used for estimation of the particle size. The obtained particle size is then used to assign them to lipoprotein subfractions obtained by other methods (e.g. gel chromatography and ultracentrifugation). The subfractions defined by the technology which we used correspond essentially to small, medium, large and very large HDL (L1–L4); very small, small, medium, large and very large LDL (L5–L9); IDL (L10); small and large VLDL (L11 and L12); remnants (L13); small and large chylomicrons (L14 and L15) (see Supplementary Material, Table S1 and (18)). In a number of studies, the 15 lipoprotein subfractions showed good agreements with results from standard enzymatic methods (16,19). However, for the calculation of particle numbers from the NMR data, one has to make additional assumptions about the shape, density and composition of these particles. Recent Cryo-EM studies showed that lipoproteins exhibit not only spherical shapes but also oval as well as almost rectangular shapes (22). However, a straightforward calculation of the volume distributions from translational diffusion coefficients is possible only for spherical geometries in the Einstein–Smoluchowski relation through the Einstein–Stokes equation. Since these additional assumptions may bias the statistical analysis, we used the concentrations c_i of methyl groups from cholesterol and fatty acids in the different particle classes L_i ($i = 1, \ldots, 15$), which can be directly measured by NMR. For assigning the lipoprotein subfractions to the serum lipids, we classified them together in a cluster analysis. To further characterize the lipoprotein subfractions, we analyzed plasma samples from 15 young men for which we had measurements at three fasting time points, as well as at seven time points during a lipid-tolerance test. Then, we calculated associations between the 15 NMR-measured lipoprotein subfractions and genetic variants within 95 lipid loci identified in GWAS. The inter-relationship among the lipoprotein subfractions and the associations of the lipid loci were analyzed in 1791 plasma samples of the population-based KORA study (23). The replication of the significant results of the lipid loci analysis was conducted in 1940 samples of the Genetic Regulation of Arterial Pressure of Humans in the Community (GRAPHIC) study (24). In the following, we refer to serum lipids as the four traits HDL-C, LDL-C, TG and TC, which were measured using standard clinical and biochemical protocols, whereas we refer to lipoprotein subfractions as L1–L15, which were measured in plasma.

Our aim is to investigate whether NMR-based fine mapping of lipoprotein subfractions provides additional information on the lipid metabolism and strengthens the associations with known genetic loci. This study provides evidence that lipoprotein subfractions provide a more in-depth view into biological processes of the lipid metabolism. This insight is complemented by a genetic characterization of the lipoprotein subfractions. In addition, we describe a lipoprotein subfraction that does not cluster together with the serum lipids. Moreover, we demonstrate that the analysis of subfractions reveals more and stronger associations with genetic loci than the analysis of serum lipids.

RESULTS

Inter-relationship of lipoprotein subfractions

In order to be independent of assumptions about the shape of lipoprotein subfractions, we assigned them to the serum lipids in a statistical analysis. First, using linear regressions with all lipoprotein subfractions as explaining variables, we observed that they explained a high proportion of serum lipid variance: 94% of the variance of TG, 84.6% of TC, 82.5% of HDL-C and 75.7% of LDL-C. To get a more in-depth view into the inter-relationship of lipoprotein subfractions, we conducted a cluster analysis of the subfractions in KORA based on their correlation matrix as a distance measure, followed by bootstrap replications to test the robustness of the clustering. The results of this cluster analysis were displayed in an unrooted tree (Fig. 1A). At first observation, the tree indicated that L1 is separate from the remaining subfractions. Furthermore, two major groups were distinguished: L2–L7 and L8–L15. Each of the two major groups contained two subgroups. In total, we had the following five clusters: (L1), (L2, L3, L5, L7), (L4, L6), (L8, L10) and (L9, L11, L12, L13, L14, L15). For the mentioned intersections, the bootstrap replications revealed an approximately unbiased (AU) probability of 1 and an s.e. of 0, which means that these divisions are absolutely reliable (Supplementary Material, Fig. S1A and Table S2). In the next step, we added the serum lipids to the tree to get a lipid-based characterization of the subfractions. After their inclusion, the main inter-relationships between the subfractions remained unchanged (Fig. 1B). We found that HDL-C clustered together with (L2, L3, L5, L7), LDL-C and TC with (L4, L6) and TG with (L9, L11, L12, L13, L14, L15). In the tree with serum lipids, the AU probabilities were smaller than before but the divisions in the mentioned five clusters were still very reliable (Supplementary Material, Fig. S1B and Table S3). We used Pearson correlations to further characterize the relations between lipoprotein subfractions and serum lipids. The results revealed that the largest correlation of HDL-C was with L3, of LDL-C and TC with L6 and of TG with L11 (Supplementary Material, Table S4). Surprisingly, lipoprotein subfraction L1 was only weakly correlated with all serum lipids.

Lipoprotein subfractions after nutritional intervention

To investigate whether the clustering of the subfractions was stable after nutritional intervention, we repeated the clustering in plasma samples from 15 young men for whom we conducted lipoprotein subfraction measurements at three fasting time points as well as at seven time points during a lipid-tolerance test. In the cluster plot of the fasting time points, we replicated the three main clusters: L1, L2–L7 and L8–L15 (Fig. 1C). For these intersections, we had reliable AU probabilities and standard errors (Supplementary Material, Fig. S1C and Table S5). In contrast to the fasting cluster plot, we observed changes in the clustering of the measurements during the lipid-tolerance test (Fig. 1D). The lipoprotein subfractions shifted and generated new groups, e.g. (L1, L6, L8). In the fasting cluster plot, L1 was independent of all other subfractions, L6 belonged to the group (L4, L6) and L8 belonged to the group (L8, L10). During the lipid-tolerance test, subfraction L13 was independent of the other subfractions. Moreover, subfraction L7 changed from the group (L2, L3, L5, L7) into the group (L7, L9, L10, L11, L12, L14, L15). For the major divisions, we observed again absolutely reliable AU probabilities and standard errors (Supplementary Material, Fig. S1D and Table S6). Subfractions within each group showed a similar trend during the lipid-tolerance test (Fig. 2). In group (L7, L9, L10, L11, L12, L14, L15), all subfractions increased after 30 min but to a different extent. L7 and L10 increased only by about 10%, whereas L14 increased by about 50%. The subfractions (L2, L3, L4, L5) stayed nearly constant during the lipid-tolerance test while lipoprotein subfraction L13 decreased notably by about 20%. Thus, the lipid-tolerance test revealed the different influences of nutritional intervention on lipoprotein subfractions. As a result, subfraction L1 was assorted together with subfractions L6 and L8, which was in contrast to the results of the fasting samples (Fig. 1D).

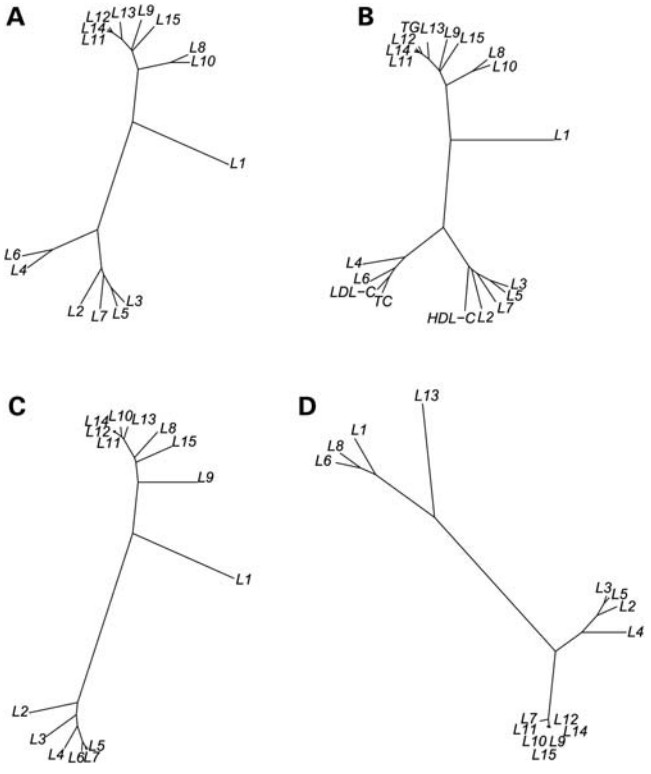

Figure 1. Cluster plots of lipoprotein subfractions. The cluster plots of the inter-relationship of the lipoprotein subfractions were displayed in an unrooted phylogeny tree using the correlation between the subfractions as distance measure. The length of a branch represents the distance between the subfractions. Each phylogeny tree was created out of 10 000 bootstrap replications. (**A**) Lipoprotein subfractions in KORA; (**B**) lipoprotein subfractions and serum lipids in KORA; (**C**) lipoprotein subfractions in the fasting samples of the 15 additional young men; and (**D**) lipoprotein subfractions in the samples during the lipid-tolerance test of the 15 young men.

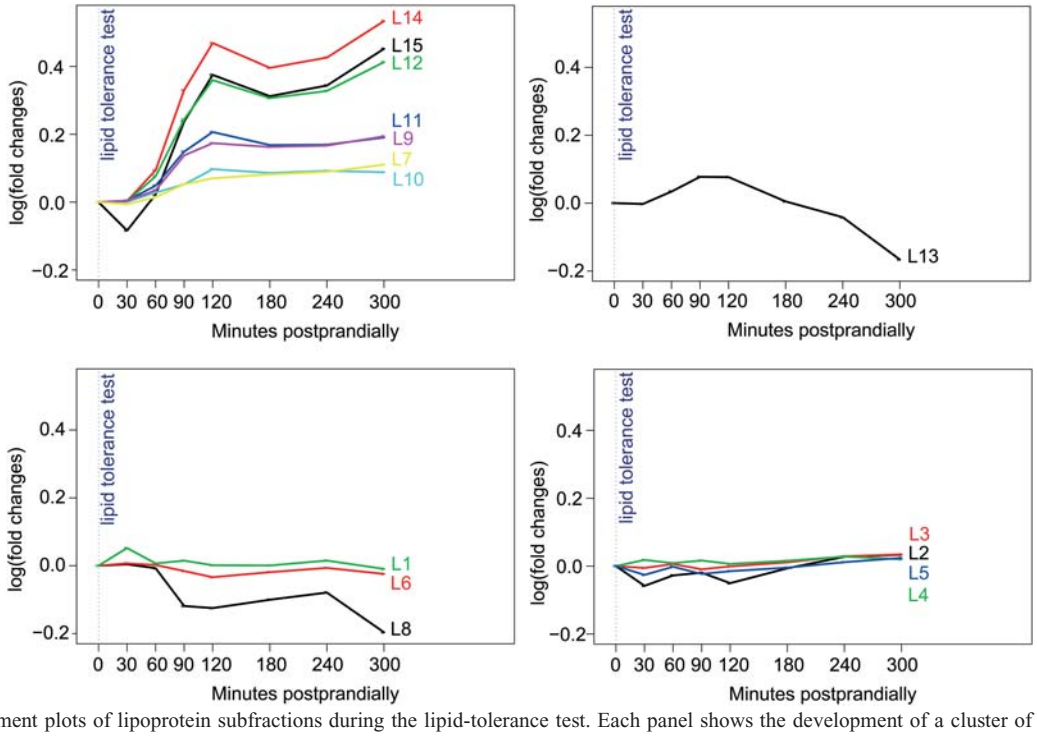

Figure 2. Development plots of lipoprotein subfractions during the lipid-tolerance test. Each panel shows the development of a cluster of the lipoprotein subfractions during the lipid-tolerance test. The *x*-axis represents the time, the *y*-axis the log-fold change, which describes the percent of change of a measurement compared with the first measurement.

Proportion of variance explained through genes increases for lipoprotein subfractions

With this mapping of the lipoprotein subfractions, we next tested the association between 101 SNPs at the 95 known serum lipid loci and the 15 subfractions using an additive genetic model. In addition to the Bonferroni-corrected significance level of 3.3×10^{-5} ($=0.05/101 \times 15$), we compared the *P*-value of the subfractions with the *P*-value of the serum lipids through calculation of a *P*-gain (see Material and Methods for details). Eight of the analyzed loci showed significant associations with at least one of the 15 subfractions (Table 1; Supplementary Material, Tables S7 and S8 and Fig. S2). In addition, associations with *FADS1-2-3*, *LIPC*, *PLTP*, *APOB* and *APOA1* had significant *P*-gains in KORA, whereas for *SORT1*, *CETP* and *GCKR*, use of NMR-measured subfractions did not strengthen the original association (Fig. 3). For *LIPC*, *CETP*, *PLTP*, *FADS1-2-3* and *GCKR*, we replicated all significant associations as well as the significant *P*-gains in the GRAPHIC Study (Supplementary Material, Table S7). For the remaining loci, some associations were not significant in GRAPHIC after Bonferroni correction. Nevertheless, the direction of effect at these loci was consistent in KORA and GRAPHIC. In contrast, when analyzing associations of serum lipids together with lipid loci, we found that only four loci in KORA were associated (*SORT1*, *CETP*, *GCKR*, *APOA1*). In addition to this, for *FADS1-2-3*, *LIPC*, *PLTP* and *APOB*, the explained variance was clearly larger for lipoprotein subfractions than for serum lipids (Fig. 3). In detail, the explained variances between lipid loci and subfractions were up to 2.3% (*APOA1* and L8). For serum lipids, we explained up to 1.7% of the variance (*CETP* and HDL-C). Altogether, the explained variance of the lipoprotein subfractions ranged from 1.5% (L9) to 4.5% (L8) and of serum lipids from 1.0% (TC) to 3.3% (TG). Summing up these results, we found more significant associations with lipoprotein subfractions and in addition, we could explain more of the variance of lipoprotein subfractions than of serum lipids. As a biological classification of the eight significant genes, Figure 4 integrates the genes together with the analyzed lipoprotein subfractions in the lipid metabolism. The colors indicate the assignment of the lipoprotein subfractions to the three main clusters: L1, L2–L7 and L8–L15.

When combining the observations made in the cluster analysis with the significant results of the association analysis, we detected comparable inter-relationships between the lipoprotein subfractions in both analyses. In the genetic analysis, we found that all lipoprotein subfractions of the cluster (L2, L3, L5, L7), which is correlated with HDL-C, were associated with *LIPC*, whereas the subfractions L2 and L3 were also associated with *CETP*. With regard to subfraction L6 of the cluster (L4, L6) together with LDL-C and TC, we found a significant association with *SORT1* (Table 1). When considering the association between L4 and *SORT1*, we saw an effect, although it was not significant (*P*-value: 3.58×10^{-5}; Table 1). The subfractions L8 and L10, which built cluster (L8, L10), were associated with *APOB*, *GCKR* and *APOA1*. Subfractions L12 and L14 and subfractions L11, L12 and L13 of cluster (L9, L11, L12, L13, L14, L15) together with TG were associated with *GCKR* and *APOA1*, respectively. Lipoprotein subfraction L1 was separate and was associated only with *PLTP* with a significant *P*-gain. Although lipoprotein subfraction

Table 1. Loci with significant associations with 15 lipoprotein subfractions

Lipoprotein subfraction	LIPC rs1532085	CETP rs3764261	PLTP rs6065906	FADS1-2–3 rs174546	SORT1 rs629301	GCKR rs1260326	APOB rs1042034	APOA1 rs964184	Max (*P*-gain)
L2	$\mathbf{3.40 \times 10^{-7}}$	$\mathbf{1.43 \times 10^{-5}}$	0.232	0.699	0.114	0.477	0.144	0.954	4.71×10^{4}
L3	$\mathbf{4.22 \times 10^{-7}}$	$\mathbf{3.59 \times 10^{-7}}$	$\mathbf{1.72 \times 10^{-5}}$	7.61×10^{-3}	0.845	0.420	0.196	0.706	3.80×10^{4}
L5	$\mathbf{5.27 \times 10^{-11}}$	7.32×10^{-4}	0.016	0.019	0.622	0.859	0.271	0.822	3.04×10^{8}
L7	$\mathbf{7.28 \times 10^{-10}}$	0.0191	0.150	0.148	0.050	0.709	0.452	0.966	2.20×10^{7}
L4	5.25×10^{-5}	5.16×10^{-3}	0.423	$\mathbf{1.38 \times 10^{-5}}$	3.58×10^{-5}	0.082	0.780	0.865	1.59×10^{3}
L6	3.33×10^{-4}	0.431	0.136	0.034	$\mathbf{1.46 \times 10^{-5}}$	1.54×10^{-3}	9.94×10^{-3}	3.04×10^{-4}	4.81×10^{1}
L1	0.015	0.584	$\mathbf{4.86 \times 10^{-7}}$	0.234	5.55×10^{-3}	0.170	0.589	0.356	1.58×10^{5}
L8	0.583	0.084	0.806	0.913	0.024	$\mathbf{9.25 \times 10^{-6}}$	$\mathbf{1.08 \times 10^{-5}}$	$\mathbf{4.82 \times 10^{-12}}$	3.42×10^{3}
L10	0.788	0.104	0.383	0.073	0.017	$\mathbf{3.73 \times 10^{-6}}$	$\mathbf{1.63 \times 10^{-5}}$	$\mathbf{9.47 \times 10^{-11}}$	2.13×10^{3}
L9	0.092	0.311	0.534	0.035	5.97×10^{-4}	4.07×10^{-3}	0.385	8.56×10^{-3}	0.622
L11	0.692	0.107	0.204	4.56×10^{-3}	7.70×10^{-3}	3.72×10^{-5}	0.025	$\mathbf{6.25 \times 10^{-7}}$	4.80
L14	0.457	0.089	0.149	1.72×10^{-3}	0.014	$\mathbf{2.01 \times 10^{-5}}$	0.196	4.49×10^{-5}	12.7
L12	0.832	0.133	0.169	5.20×10^{-3}	0.014	$\mathbf{6.88 \times 10^{-6}}$	6.04×10^{-3}	$\mathbf{2.72 \times 10^{-7}}$	5.75
L13	0.199	0.191	0.316	2.23×10^{-3}	0.032	1.84×10^{-4}	0.087	$\mathbf{3.16 \times 10^{-7}}$	9.80
L15	0.419	0.607	0.493	2.73×10^{-3}	0.067	9.85×10^{-4}	0.731	4.56×10^{-4}	8.01
max(*P*-gain)	3.04×10^{8}	0.015	1.58×10^{5}	1.59×10^{3}	0.628	0.014	3.21×10^{3}	3.42×10^{3}	
KORA HDL-C	0.016	$\mathbf{5.69 \times 10^{-9}}$	0.153	0.907	0.173	0.525	0.257	0.027	
KORA LDL-C	0.947	0.341	0.759^{1}	0.594	$\mathbf{1.36 \times 10^{-5}}$	0.214	0.088	0.718	
KORA TC	0.220	0.259	0.444	0.416	$\mathbf{8.71 \times 10^{-6}}$	0.020	0.056	0.069	
KORA TG	0.212	0.105	0.077	0.022	0.012	$\mathbf{2.54 \times 10^{-7}}$	0.035	$\mathbf{2.09 \times 10^{-8}}$	
GLC HDL-C	$\mathbf{2.92 \times 10^{-96}}$	$\mathbf{7.10 \times 10^{-380}}$	$\mathbf{1.90 \times 10^{-22}}$	$\mathbf{2.62 \times 10^{-22}}$	6.19×10^{-8}	0.078	$\mathbf{1.22 \times 10^{-30}}$	$\mathbf{5.21 \times 10^{-47}}$	
GLC LDL-C	0.852	$\mathbf{1.64 \times 10^{-12}}$	0.297	$\mathbf{1.76 \times 10^{-21}}$	$\mathbf{9.70 \times 10^{-171}}$	2.33×10^{-4}	$\mathbf{8.32 \times 10^{-25}}$	$\mathbf{1.47 \times 10^{-26}}$	
GLC TC	$\mathbf{8.83 \times 10^{-20}}$	$\mathbf{6.67 \times 10^{-14}}$	0.970	$\mathbf{2.85 \times 10^{-22}}$	$\mathbf{5.77 \times 10^{-131}}$	$\mathbf{7.31 \times 10^{-27}}$	$\mathbf{3.71 \times 10^{-18}}$	$\mathbf{6.21 \times 10^{-57}}$	
GLC TG	$\mathbf{1.78 \times 10^{-11}}$	$\mathbf{6.15 \times 10^{-12}}$	$\mathbf{2.59 \times 10^{-17}}$	$\mathbf{5.41 \times 10^{-24}}$	0.062	$\mathbf{5.68 \times 10^{-133}}$	$\mathbf{1.36 \times 10^{-45}}$	$\mathbf{6.71 \times 10^{-240}}$	

This table shows the *P*-values of the eight loci which were associated with at least one of the 15 lipoprotein subfractions. Results were provided for the 15 lipoprotein subfractions (L1–L15), serum lipids in KORA (KORA HDL-C, KORA LDL-C, KORA TC, KORA TG) and the global lipids' meta-analysis (http://www.sph.umich.edu/csg/abecasis/public/lipids2010/, last accessed on 14 December 2011) (GLC HDL-C, GLC LDL-C, GLC TC, GLC TG) (1). The lipoprotein subfractions were ordered according to the hierarchical clustering of Figure 3. *P*-values highlighted in bold were significant after Bonferroni correction. The *P*-values for the associations between the subfractions and the eight loci were visualized in Supplementary Material, Figure S2. Detailed results for the significant associations (bold) were summarized in Supplementary Material, Table S7.

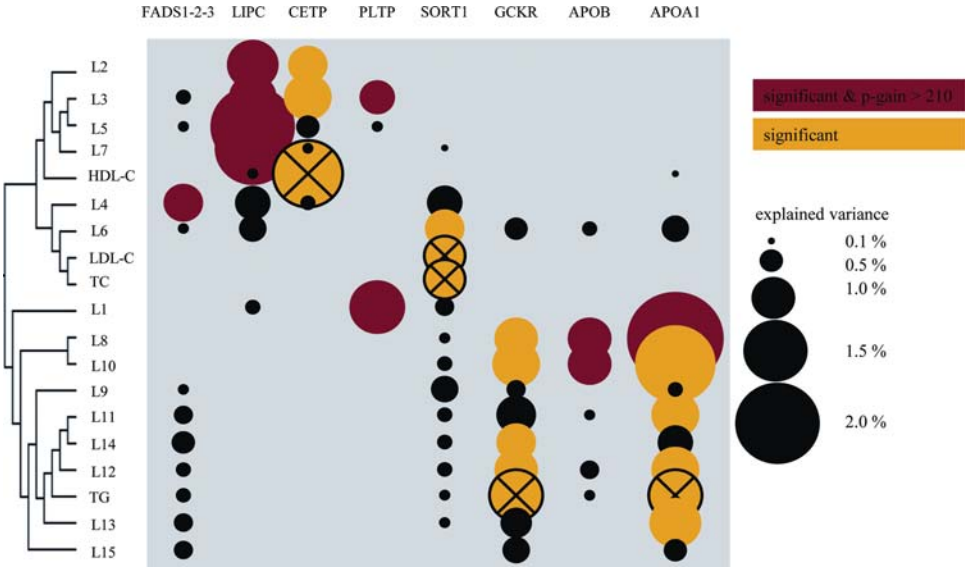

Figure 3. Explained variance of the associated loci. This figure presents the explained variances shown for the lipoprotein subfractions having a *P*-value of <0.05 for association with lipid loci. The diameter of each circle represents the explained variance, a circle highlighted in yellow corresponds to a significant association and a circle colored in red corresponds to a significant association with significant *P*-gain. Circles with a black cross belong to serum lipids. The lipoprotein subfractions were ordered according to a hierarchical clustering which is displayed on the *y*-axis of this figure.

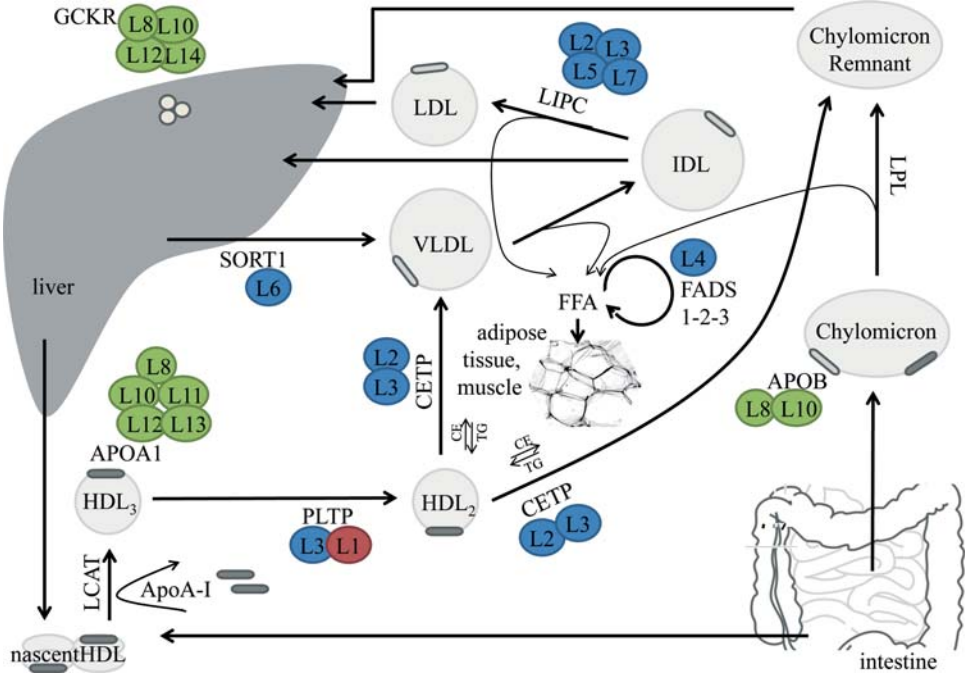

Figure 4. Classification of lipoprotein subfractions and associated loci in the lipid metabolism. This figure combines the results of our association analyses with the lipid metabolism. We displayed each associated gene at least once in this figure and attached the associated lipoprotein subfractions to them. For clarity, we restricted the lipid metabolism to pathways where our associated loci are involved. The color of the lipoprotein subfractions encodes the membership to a cluster. We assigned the lipoprotein subfractions to the three larger clusters L1, L2–L7 and L8–L15, to keep the figure clear.

L3 was also associated with *PLTP*, the effect was in opposite directions for L1 and L3 (Supplementary Material, Table S7). In conclusion, the genetic analysis confirms the observations made in the clustering and reveals further information about biological aspects of the lipoprotein subfractions.

DISCUSSION

Clustering reveals that L1 is not captured by serum lipids

Clustering of the lipoprotein subfractions measured in fasting samples together with HDL-C, LDL-C, TG and TC revealed

five groups of subfractions. HDL-C clustered together with L2, L3, L5, and L7, whereas LDL-C and TC clustered together with L4 and L6 and TG clustered together with L9, L11, L12, L13, L14 and L15. In addition, we detected that lipoprotein subfraction L1 does not cluster together with the serum lipids. Due to its size, L1 is considered to correspond to the smallest HDL subfraction. This finding matches the observations made by others that the smallest HDL subfraction behaves in a different way than the larger HDL subfractions (11,25). Inouye *et al.* (25) speculated that the smallest HDL subfraction may have pro-atherogenic potential which is in contrast to the anti-atherogenic properties of HDL-C. However, conflicting data on the association between cardiovascular disease risk and small HDL fractions as assessed by different techniques still complicate painting a concise picture of the fractions' specific role (26). HDL-C clustered together with L2, L3, L5 and L7, which are considered to correspond to medium and large HDL as well as very small and medium LDL, respectively. Interestingly, in addition to HDL-related subfractions, LDL-related subfractions also clustered together with HDL-C. Furthermore, LDL-C clustered together with L4 and L6, which is considered to be related to very large HDL and small LDL, respectively. This cross-mixed correlation of HDL and LDL subfractions needs further investigation. The subfractions clustered together with TG are related to the more TG-rich subfractions of VLDL and chylomicrons. When clustering the lipoprotein subfractions measured in plasma taken during a lipid-tolerance test, we got different groups of subfractions. The analysis of the lipoprotein subfractions during the lipid-tolerance test revealed that some subfractions were increased in response to a standardized lipid-tolerance test, whereas other subfractions stayed nearly constant. While subfractions which cluster together with TG tend to increase after nutritional intervention, subfractions which cluster together with HDL-C stay the same. Interestingly, subfraction L13, which relates to remnants, behaves different than the other subfractions which cluster together with TG. Thus, nutritional intervention had different influences on distinct lipoprotein subfractions. The analysis of samples during the lipid-tolerance test was carried out only in 15 subjects. However, this trial was highly controlled and clustering of the lipoprotein subfractions at fasting time points led to a clustering comparable with that of KORA samples.

Using lipoprotein subfractions, we identified eight loci that were significantly associated in the KORA study, whereas when analyzing TC, HDL-C, LDL-C and TG in the same individuals, we found only half of the loci. These eight loci contribute to diverse mechanisms of the lipid metabolism such as regulatory elements or structural lipid components which is illustrated in Figure 4.

PLTP indicates the role of L1 in the lipid metabolism

PLTP encodes for the phospholipid transfer protein that transfers phospholipids and other amphipathic compounds between lipoprotein particles (4,27) (Fig. 4). Although the role of the phospholipid transfer protein in the reverse cholesterol transport has long been studied, it still remains controversial (28). It has been shown in a large meta-analysis on serum lipids that *PLTP* is significantly associated with HDL-C and TG levels (1) as well as with HDL particle size (11,20). Our analysis revealed that notably the lipoprotein subfraction L1, which was only weakly correlated with HDL-C, and lipoprotein subfraction L3 were associated with *PLTP* with opposite directions of effects. The other subfractions L2, L5 and L7 which clustered together with HDL-C showed no association. Here, the subfractions reveal an in-depth insight into the lipid metabolism. The opposite directions of effect of the association of L1 and L3 presumably compensate each other partly when analyzing serum HDL-C. Moreover, due to the opposite directions of effect, it can be speculated that *PLTP* is involved in the conversion of L1 to L3. In addition, lipoprotein subfraction L1 was only marginally captured by the measurements of serum lipids as L1 was weakly negatively correlated with HDL-C and weakly positively with the other serum lipids. Therefore, it is possible that L1 is involved in parts of the lipid metabolism which were not covered by the measurement of HDL-C. As L1 is related to the smallest HDL subfraction, it is assumed that L1 represents nascent HDL which would be an explanation for a negative correlation with HDL-C.

Lipoprotein subfractions revealed in-depth insight into mechanisms of *LIPC*, *CETP* and *FADS1-2-3*

LIPC encodes for hepatic lipase which catabolizes TG-enriched HDL (4) and breaks down TG to diacyl- and monoacylglycerols and fatty acids. This molecular function is observed in associations between *LIPC* and numerous concentrations of glycerophosphatidylcholines, glycerophosphatidylethanolamines and sphingomyelins (29). In our analysis, the strongest association occurred with L5 and L7, which clustered together with HDL-C and are considered to be related to very small and medium LDL, respectively. Here, we observed the largest increase in the proportion of explained variance when compared with serum lipids. But also L2 and L3, the other lipoprotein subfractions which clustered together with HDL-C, were associated with *LIPC*. Interestingly, although all subfractions which cluster together with HDL-C were significantly associated with *LIPC* with the same direction of effect, the association between *LIPC* and HDL-C itself was not significant (P-value 1.60×10^{-2}). For the remaining subfractions, especially for the subfractions correlated with TG, we did not see an association with *LIPC* as it is observed by others (11). Although *LIPC* is associated with all four lipoprotein subfractions that cluster together with HDL-C, *CETP* was only associated with L2 and L3. *CETP* encodes a protein that exchanges cholesteryl esters for TG between lipoproteins (30) (Fig. 4). The *FADS1-2-3* gene complex encodes for key enzymes in the metabolism of long-chain polyunsaturated fatty acids. Our analyses revealed associations between *FADS1-2-3* and L4, an LDL-C-correlated subfraction which is considered to be related to large HDL. For LDL-C itself, we did not see an association with *FADS1-2-3*. Although *FADS1-2-3* is strongly associated with TG in the global lipids meta-analysis in more than 100 000 samples (1), we observed only a small effect which was not significant when based on the analysis of 1791 samples. The strong association between *FADS1-2-3* and L4 highlighted the potential of lipoprotein subfractions and hinted at further biological

implications of the *FADS1-2-3* gene complex in the lipid metabolism.

More insight in pathway regulation and genes which encode structural components

Among others, *SORT1* and *GCKR* are genes that are involved in pathways regulating lipid and glucose metabolism. Musunuru *et al.* (31) showed that hepatic expression of *SORT1* alters LDL-C and VLDL levels and that *SORT1* is associated with coronary artery disease. In more detail, *SORT1* encodes sortilin which presumably controls the biogenesis and hepatic release of VLDL from which LDL is generated by lipolysis (32) (Fig. 4). In our analysis, *SORT1* was associated with L6, which clustered together with LDL-C and relates to small LDL. *APOB* and *APOA1* are genes that encode the structural components apolipoprotein B and apolipoprotein A-I. Apolipoprotein B is the main apolipoprotein of chylomicrons, VLDL, IDL, LDL and Lp(a), whereas apolipoprotein A-I is the main apolipoprotein of HDL (4,33) (Fig. 4). In our analysis of KORA samples, both genes were predominantly associated with lipoprotein subfractions L8 and L10. These subfractions did not cluster closely with one of the serum lipids but were more related to the TG-correlated subfractions L9, L11, L12, L13, L14 and L15. These subfractions relate to VLDL as well as chylomicron subfractions. While *APOB* was associated only with L8 and L10, *APOA1* also showed associations with the particles L11, L12 and L13 in KORA. The associations of *APOA1* and *APOB* with L8 had the same direction of effect in KORA and GRAPHIC samples although we did not significantly replicate them.

CONCLUSION

We showed that lipoprotein subfractions provide a more detailed insight into the lipid metabolism and thus strengthen the association with disease-relevant genetic loci. Chasman *et al.* (11) reported 43 loci associated with lipoprotein subfractions when analyzing 17 296 women. At that time, 10 of these loci were novel findings. By now, some of these loci were also found by Teslovich *et al.* (1) in a serum lipid meta-analysis of more than 100 000 samples. Kaess *et al.* (20) observed a strengthening in association when analyzing HDL size and HDL particle number. In our results, we observed an increase in the proportion of variance explained when analyzing lipoprotein subfractions instead of serum lipids. With the eight loci, we explained up to 4.5% of the variance of the lipoprotein subfractions, whereas only up to 3.3% of the variance of serum lipids could be explained.

In this study, we showed that analyzing well-defined lipoprotein subfractions together with known genetic lipid loci leads to a genetic characterization of the lipoprotein subfractions as well as an in-depth insight into various processes of the lipid metabolism. We identified five distinct groups of lipoprotein subfractions, one of them (L1) was only marginally captured by serum lipids and therefore extends our knowledge of lipoprotein biochemistry. During a lipid-tolerance test, the relationship between the individual classes changed and L1 lost its special position. Based on this initial specification of the lipoprotein subfractions, further testing in clinical samples will reveal more information on their biological nature and their impact in disease-causing mechanisms. In conclusion, NMR-based fine mapping of lipoprotein subfractions provides novel information on their biological nature and strengthens the association with genetic loci.

MATERIAL AND METHODS

Study description

KORA study. The Cooperative Health Research in the Region of Augsburg (KORA) study is a series of independent population-based epidemiological surveys and follow-up studies of participants living in the region of Augsburg, Southern Germany (23). All participants are residents of Germany with a German nationality identified through registration. All participants gave signed informed consent. The local ethics committee has approved the studies. The present study includes data of the follow-up study KORA F4 (2006–2008) of the KORA S4 survey (1999/2000). For genotyping, we included 1814 randomly selected participants of KORA F4. The KORA F4 samples were genotyped with the Affymetrix Human SNP Array 6.0 and imputed with IMPUTE v.0.4.2 based on Hap Map II (34).

Blood collection. We collected blood samples between 2006 and 2008 during the KORA F4 examinations. To avoid variation due to circadian rhythm, blood was drawn in the morning between 8:00 a.m. and 10:00 a.m. after a period of overnight fasting. One part of the blood was drawn into serum gel tubes, gently inverted two times and then allowed to rest for 30 min at room temperature (18 − 25°C) to obtain complete coagulation. The material was then centrifuged at 15°C for 10 min at 2750*g*. Serum was divided into aliquots and kept for a maximum of 6 h at 4°C, after which it was deep-frozen to −80°C until analysis. Another part of the blood was drawn into ethylene diaminetetraacetic acid (EDTA) tubes, gently inverted two times and left on the Sarstedt Universal mixer less than 5 min to avoid mechanical hemolysis, followed by centrifugation at 15°C for 10 min at 2750*g*. Thereafter, plasma was separated, divided into 200 μl aliquots and kept at 4°C, after which it was deep-frozen to −80°C. After less than 2 weeks, plasma was stored in the gaseous phase of liquid nitrogen (−196°C). Following the transport on dry ice to Regensburg, it was deep-frozen at −80°C for 2 months. Then, plasma was thawed and immediately analyzed.

Serum lipid measurements. All serum lipids were measured on fresh samples using the Dimension RxL (Dade Behring). TC was determined by cholesterol esterase method (CHOL Flex, Dade-Behring, CHOD-PAP method), HDL-C using the AHDL Flex (Dade-Behring, CHOD-PAP method after selective release of HDL-C), LDL-C using the ALDL Flex (Dade Behring, CHOD-PAP method after colorless usage of all non-LDL-C) and TG was measured using a TGL Flex (Dade Behring, enzymatic colorimetric test, GPO-PAP method). For the analysis, all lipid values were naturally log-transformed to achieve normality. Summary statistics for

serum lipids and lipoprotein subfractions are combined in Supplementary Material, Table S9.

GRAPHIC study. The GRAPHIC study recruited 2024 individuals from 520 nuclear families of white European origin from Leicestershire in the UK (24). The details of recruitment, phenotyping and sample analysis have been reported elsewhere (24). In brief, for families to be included, both parents had to be aged 40–60 with two offspring aged 18 or over, with all members agreeing to take part in the study. A standardized questionnaire was used to obtain a comprehensive medical history from participants followed by physical examination, anthropometric measurements, clinic and 24 h ambulatory blood pressure monitoring. The standard biochemistry measurements including TC and HDL-C were performed on non-fasting serum samples using enzymatic assays in an Olympus AU5430 analyzer (35). Summary statistics for TC, HDL-C and lipoprotein subfractions are combined in Supplementary Material, Table S9.

Trial of lipid-tolerance test. Fifteen young and metabolically healthy men were recruited with a very narrow age range (22–33 years) and normal body mass index (20–25 kg/m^2) to undergo a lipid-tolerance test. For comparison, fasting samples were taken on three days at 8:00 a.m. The second fasting sample was taken 4 weeks after the first fasting sample. The third sample was taken 24 h later. This trial was approved by the ethical commission of the Technische Universität München (no. 2087/08).

Plasma collection. Blood samples were collected into 9 ml EDTA K_2-Gel tubes (Sarstedt, Nümbrecht, Germany). EDTA-tubes were immediately centrifuged at 3000*g* for 10 min at 20°C. Plasma was aliquoted by an automatic pipette and was immediately deep-frosted on dry ice and stored at −80°C until analysis, except for the duration of the transport to Regensburg on dry ice.

Oral lipid-tolerance test. The oral lipid-tolerance test drink consisted of a 3:1 mixture, containing three parts Fresubin® Energy Drink chocolate (Fresenius Kabi, Bad Homburg, Germany) and one part Calogen® (Nutricia, Zoetemeer, Netherlands). Calogen® is a fat emulsion containing 50 g of long-chain TG per 100 ml. The test drink was served at room temperature at 8:00 a.m. after an overnight fast for ingestion within 5 min. Plasma collections were performed after 0, 30, 60, 90, 120, 180, 240 and 300 min after the lipid ingestion. Summary statistics for lipoprotein subfractions are combined in Supplementary Material, Table S9.

Lipoprotein subfraction analysis. The lipoprotein subfraction distribution was assessed by NMR spectroscopy and carried out at LipoFIT GmbH (Regensburg, Germany). The technology has been patented (US7927878; AU2005250571; DE 10 2004 026 903 B4) (36–38). Briefly, diffusion-weighted NMR spectra of blood plasma were recorded on a Bruker 600 MHz spectrometer Avance II^{plus}, which revealed characteristic overall profiles of the lipoprotein signals. Using the LipoFIT proprietary software, the spectral regions of the spectra ranging from 0.6 to 1.5 ppm were decomposed into a set of 15 lipoprotein classes termed L1–L15 that are characterized by different diffusion constants. The classes were defined in such a way that the corresponding diffusion constants agreed with the presumed particle sizes given in Supplementary Material, Table S1. A direct outcome of the analysis of the methyl resonance is the concentration of methyl groups from cholesterol and fatty acids c_i in each NMR-defined lipoprotein subfraction i. For the statistical analysis, the natural logarithm $\ln c_i$ of the concentration was used.

Statistical analyses

Association with 95 lipid loci. For a genetic characterization of the measured lipoprotein subfractions, we analyzed in KORA the 101 candidate SNPs on 95 lipid loci described by Teslovich *et al.* (1). We used the software QUICKTEST with an additive model to calculate associations. To correct for multiple testing, we applied Bonferroni correction for the 101 candidate SNPs and 15 lipoprotein subfractions, i.e. $P\text{-value} < 0.05/101 \times 15$. To test if we had an increase in information due to analyzing lipoprotein subfractions compared with serum lipids, we calculated a *P*-gain defined as.

$$P\text{-gain(lipoprotein subfraction)} = \frac{\min(P\text{-value(HDL-C)}, P\text{-value(LDL-C)}, P\text{-value(TG)}, P\text{-value(TC)})}{P\text{-value(lipoprotein subfraction)}}.$$

To define critical values for the *P*-gain, we derived the distribution of a universal *P*-gain, i.e. for the uncorrelated phenotypes P_1 and P_2 with $P\text{-gain}(P_2) = P\text{-value}(P_1)/P\text{-value}(P_2)$, which is conservative to the *P*-gain which we used here (A.K. Petersen, J. Krumsiek, B. Wägele, F. Theis, H.E. Wichmann, C. Gieger and K. Suhre, manuscript in preparation). For this conservative *P*-gain, the critical value of 10 corresponds to a *P*-value of 0.05 and Bonferroni correction for multiple testing can be done by multiplying the critical value 10 with the number of loci to be corrected. As we detected 21 significant associations in the discovery step, we set the critical value to 210. For replication, we applied multiple testing for 10 loci, which lead to a critical value of 100. A comparable gain was introduced by Altmaier *et al.* (39) and Gieger *et al.* (29) and applied by Illig *et al.* (34) and Suhre *et al.* (40). The explained variance of an SNP was calculated as the difference between explained variance of a linear model with SNP, age and sex as explaining variables and of a linear model with only age and sex as explaining variables.

Replication. *In silico* replication of the significant associations in the KORA Study was conducted in the GRAPHIC Study using information from the Illumina HumanCVD BeadChip array (24). The analysis of association was carried out using generalized estimation equations (with exchangeable correlation structure to account for familial correlations) with adjustment for age, age^2 and sex under additive model of inheritance (24). To correct for multiple testing, we applied a Bonferroni correction for the 21 significant SNP–lipoprotein subfraction associations.

Cluster dendrogram. For the visualization of the correlation structure within the lipoprotein subfraction dataset, we used an unrooted phylogeny tree where the length of each branch represents the distance between variables. We plotted the tree using the package 'ape' (41) within the R-Project environment (42). The distance measure is based on the correlation between two variables and for the clustering of the lipoprotein subfractions, the average linkage method was used. To test the robustness of our trees, we applied a bootstrap method implemented in the 'pvclust' package (43) of the R-Project with 10 000 bootstrap replications. To measure the confidence of each branch, we used the AU probability, which is more accurate than the bootstrap probability (44). The AU probability is calculated on multi-scale bootstrap resamplings. Besides AU probabilities, we also calculated standard errors to evaluate the confidence of each branch. High AU probabilities and low standard errors indicate a strong support for a branch. For the additional 15 young men, we had multiple measurements at fasting time points as well as during a lipid-tolerance test. Aiming at illustrating the variation between variables and not within individuals for the fasting dendrogram, the clustering of the lipoprotein subfractions was based on average values of multiple measurements from a participant. For the cluster plot of the lipoprotein subfractions during the lipid-tolerance test, we aimed to illustrate the variation over the time, so we calculated average values of the measurements retained at one time point from all participants. To classify the lipoprotein subfractions in a natural way, we incorporated in a second step the serum lipids in the cluster analysis of KORA samples.

Correlation matrix. To calculate the Pearson correlation matrix of lipoprotein subfractions and serum lipids, we used the 'cor' function implemented in the R-Project Environment for all pair-wise complete observations. To calculate the explained variance, we conducted a linear regression analysis for each serum lipid separately with all lipoprotein subfractions as well as age and sex as explaining variables.

Development plots. To visualize the development of the lipoprotein subfractions during the lipid-tolerance test, we plotted time-dependent graphs for each cluster discovered in the cluster dendrogram separately. To better visualize the change of the subfractions in comparison with the measurement at the starting time point, we used log-fold changes. A fold change is the ratio of a measurement at a certain time point to the measurement at the starting time point. Through calculation of the logarithm, the *y*-axis represents the change in percent with positive values as increase and negative values as decrease.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

FUNDING

The KORA research platform was initiated and financed by the Helmholtz Center Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2, NGFNPlus 01GS0823, NGFNPlus 01GS0832 and NGFNPlus 01GS0834) and through additional funds from the University of Ulm and the University Hospital Regensburg. Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ and by a grant from the BMBF to the German Center for Diabetes Research (DZD e.V.), as well as from the BMBF-funded German Network for Mitochondrial Disorders (mitoNET 01GM0862). Furthermore, the study received funding from the European Community's Seventh Framework Program (FP7/2007–2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413. In addition, the KORA study as well as the trial of the lipid-tolerance test were funded by the BMBF project Systems Biology of Metabotypes (SysMB0 0315494A and 0315494D). A.-K.P is supported by the ENGAGE Exchange and Mobility Program (HEALTH-F4-2007-201413). J.R. is supported by DFG Graduiertenkolleg 'GRK 1563, Regulation and Evolution of Cellular Systems' (RECESS) and W.R.-M. by BMBF grant 03IS2061B (project Gani_Med). Recruitment and genotyping of the Genetic Regulation of Arterial Pressure of Humans in the Community cohort was funded by the British Heart Foundation. N.J.S. holds a British Heart Foundation Chair of Cardiology. This study is part of the research portfolio supported by the Leicester National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We express our appreciation to all study participants.

Conflict of Interest statement. W.K. and H.R.K. are employed by the University of Regensburg and are co-founders of LipoFIT GmbH; F.H. and V.P. are co-founders and CEOs of LipoFIT GmbH. However, direct financial benefits did not follow from this work. F.H., W.K., H.R.K. and V.P. contributed only to logistics, optimizing NMR spectroscopy and NMR data interpretation. They were not involved in the design of the study, statistical analyses or interpretation of the results. The involvement of LipoFIT GmbH and the affiliation of F.H., W.K., H.R.K. and V.P. to LipoFIT GmbH does not alter the authors' adherence to all the Human Molecular Genetics policies in sharing data and materials.

REFERENCES

1. 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. *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707–713.
2. Castelli, W.P. (1996) Lipids, risk factors and ischaemic heart disease. *Atherosclerosis*, **124** (Suppl.), S1–S9.
3. Castelli, W.P., Doyle, J.T., Gordon, T., Hames, C.G., Hjortland, M.C., Hulley, S.B., Kagan, A. and Zukel, W.J. (1977) HDL cholesterol and other lipids in coronary heart disease. The cooperative lipoprotein phenotyping study. *Circulation*, **55**, 767–772.
4. Rader, D.J. (2006) Molecular regulation of HDL metabolism and function: implications for novel therapies. *J. Clin. Invest.*, **116**, 3090–3100.

5. Asztalos, B.F., Cupples, L.A., Demissie, S., Horvath, K.V., Cox, C.E., Batista, M.C. and Schaefer, E.J. (2004) High-density lipoprotein subpopulation profile and coronary heart disease prevalence in male participants of the Framingham Offspring Study. *Arterioscler. Thromb. Vasc. Biol.*, **24**, 2181–2187.
6. Rader, D.J. (2009) Lecithin:cholesterol acyltransferase and atherosclerosis: another high-density lipoprotein story that doesn't quite follow the script. *Circulation*, **120**, 549–552.
7. von Eckardstein, A. (2010) Implications of torcetrapib failure for the future of HDL therapy: is HDL-cholesterol the right target? *Expert Rev. Cardiovasc. Ther.*, **8**, 345–58.
8. Ala-Korpela, M., Korhonen, A., Keisala, J., Horkko, S., Korpi, P., Ingman, L.P., Jokisaari, J., Savolainen, M.J. and Kesaniemi, Y.A. (1994) 1H NMR-based absolute quantitation of human lipoproteins and their lipid contents directly from plasma. *J. Lipid Res.*, **35**, 2292–2304.
9. Otvos, J.D., Jeyarajah, E.J., Bennett, D.W. and Krauss, R.M. (1992) Development of a proton nuclear magnetic resonance spectroscopic method for determining plasma lipoprotein concentrations and subspecies distributions from a single, rapid measurement. *Clin. Chem.*, **38**, 1632–1638.
10. Festa, A., Williams, K., Hanley, A.J., Otvos, J.D., Goff, D.C., Wagenknecht, L.E. and Haffner, S.M. (2005) Nuclear magnetic resonance lipoprotein abnormalities in prediabetic subjects in the Insulin Resistance Atherosclerosis Study. *Circulation*, **111**, 3465–3472.
11. Chasman, D.I., Pare, G., Mora, S., Hopewell, J.C., Peloso, G., Clarke, R., Cupples, L.A., Hamsten, A., Kathiresan, S., Malarstig, A. *et al.* (2009) Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet.*, **5**, e1000730.
12. Decewicz, D.J., Neatrour, D.M., Burke, A., Haberkorn, M.J., Patney, H.L., Vernalis, M.N. and Ellsworth, D.L. (2009) Effects of cardiovascular lifestyle change on lipoprotein subclass profiles defined by nuclear magnetic resonance spectroscopy. *Lipids Health Dis.*, **8**, 26.
13. Irving, B.A., Nair, K.S. and Srinivasan, M. (2011) Effects of insulin sensitivity, body composition, and fitness on lipoprotein particle sizes and concentrations determined by nuclear magnetic resonance. *J. Clin. Endocrinol. Metab.*, **96**, E713–E718.
14. Ala-Korpela, M., Lankinen, N., Salminen, A., Suna, T., Soininen, P., Laatikainen, R., Ingman, P., Jauhiainen, M., Taskinen, M.R., Heberger, K. *et al.* (2007) The inherent accuracy of 1H NMR spectroscopy to quantify plasma lipoproteins is subclass dependent. *Atherosclerosis*, **190**, 352–358.
15. Ala-Korpela, M. (2008) Critical evaluation of 1H NMR metabonomics of serum as a methodology for disease risk assessment and diagnostics. *Clin. Chem. Lab. Med.*, **46**, 27–42.
16. Soutschek, J., Akinc, A., Bramlage, B., Charisse, K., Constien, R., Donoghue, M., Elbashir, S., Geick, A., Hadwiger, P., Harborth, J. *et al.* (2004) Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature*, **432**, 173–178.
17. Zimmermann, T.S., Lee, A.C., Akinc, A., Bramlage, B., Bumcrot, D., Fedoruk, M.N., Harborth, J., Heyes, J.A., Jeffs, L.B., John, M. *et al.* (2006) RNAi-mediated gene silencing in non-human primates. *Nature*, **441**, 111–114.
18. Linsel-Nitschke, P., Jansen, H., Aherrarhou, Z., Belz, S., Mayer, B., Lieb, W., Huber, F., Kremer, W., Kalbitzer, H.R., Erdmann, J. *et al.* (2009) Macrophage cholesterol efflux correlates with lipoprotein subclass distribution and risk of obstructive coronary artery disease in patients undergoing coronary angiography. *Lipids Health Dis.*, **8**, 14.
19. Kaess, B., Fischer, M., Baessler, A., Stark, K., Huber, F., Kremer, W., Kalbitzer, H.R., Schunkert, H., Riegger, G. and Hengstenberg, C. (2008) The lipoprotein subfraction profile: heritability and identification of quantitative trait loci. *J. Lipid Res.*, **49**, 715–723.
20. Kaess, B.M., Tomaszewski, M., Braund, P.S., Stark, K., Rafelt, S., Fischer, M., Hardwick, R., Nelson, C.P., Debiec, R., Huber, F. *et al.* (2011) Large-scale candidate gene analysis of HDL particle features. *PLoS One*, **6**, e14529.
21. Rubio-Aliaga, I., de Roos, B., Duthie, S.J., Crosley, L.K., Mayer, C., Horgan, G., Colquhoun, I.J., Le Gall, G., Huber, F., Kremer, W. *et al.* (2011) Metabolomics of prolonged fasting in humans reveals new catabolic markers. *Metabolomics*, **7**, 375–387.
22. Van Antwerpen, R., La Belle, M., Navratilova, E. and Krauss, R.M. (1999) Structural heterogeneity of apoB-containing serum lipoproteins visualized using cryo-electron microscopy. *J. Lipid Res.*, **40**, 1827–1836.
23. Wichmann, H.E., Gieger, C. and Illig, T. (2005) KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen*, **67** (Suppl. 1), S26–S30.
24. Tomaszewski, M., Debiec, R., Braund, P.S., Nelson, C.P., Hardwick, R., Christofidou, P., Denniff, M., Codd, V., Rafelt, S., van der Harst, P. *et al.* (2010) Genetic architecture of ambulatory blood pressure in the general population: insights from cardiovascular gene-centric array. *Hypertension*, **56**, 1069–1076.
25. Inouye, M., Kettunen, J., Soininen, P., Silander, K., Ripatti, S., Kumpula, L.S., Hamalainen, E., Jousilahti, P., Kangas, A.J., Mannisto, S. *et al.* (2010) Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol. Syst. Biol.*, **6**, 441.
26. Camont, L., Chapman, M.J. and Kontush, A. (2011) Biological activities of HDL subpopulations and their relevance to cardiovascular disease. *Trends Mol. Med.*, **17**, 594–603.
27. Huuskonen, J., Olkkonen, V.M., Jauhiainen, M. and Ehnholm, C. (2001) The impact of phospholipid transfer protein (PLTP) on HDL metabolism. *Atherosclerosis*, **155**, 269–281.
28. Yazdanyar, A., Yeang, C. and Jiang, X.C. (2011) Role of phospholipid transfer protein in high-density lipoprotein- mediated reverse cholesterol transport. *Curr. Atheroscler. Rep.*, **13**, 242–248.
29. Gieger, C., Geistlinger, L., Altmaier, E., Hrabe de Angelis, M., Kronenberg, F., Meitinger, T., Mewes, H.W., Wichmann, H.E., Weinberger, K.M., Adamski, J. *et al.* (2008) Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet.*, **4**, e1000282.
30. Boes, E., Coassin, S., Kollerits, B., Heid, I.M. and Kronenberg, F. (2009) Genetic-epidemiological evidence on genes associated with HDL cholesterol levels: a systematic in-depth review. *Exp. Gerontol.*, **44**, 136–60.
31. Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., Li, H., Kuperwasser, N., Ruda, V.M. *et al.* (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*, **466**, 714–719.
32. Kjolby, M., Andersen, O.M., Breiderhoff, T., Fjorback, A.W., Pedersen, K.M., Madsen, P., Jansen, P., Heeren, J., Willnow, T.E. and Nykjaer, A. (2010) Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metab.*, **12**, 213–223.
33. Kane, J.P., Hardman, D.A. and Paulus, H.E. (1980) Heterogeneity of apolipoprotein B: isolation of a new species from human chylomicrons. *Proc. Natl Acad. Sci.*, **77**, 2465–2469.
34. Illig, T., Gieger, C., Zhai, G., Romisch-Margl, W., Wang-Sattler, R., Prehn, C., Altmaier, E., Kastenmuller, G., Kato, B.S., Mewes, H.W. *et al.* (2010) A genome-wide perspective of genetic variation in human metabolism. *Nat. Genet.*, **42**, 137–141.
35. Samani, N.J., Braund, P.S., Erdmann, J., Gotz, A., Tomaszewski, M., Linsel-Nitschke, P., Hajat, C., Mangino, M., Hengstenberg, C., Stark, K. *et al.* (2008) The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol. *J. Mol. Med.*, **86**, 1233–1241.
36. Huber, F., Kalbitzer, H.R. and Kremer, W. (2005) *Verfahren zur Bestimmung von Lipoproteinen in Körperflüssigkeiten und Messanordnung dafür*. DE 10 2004 026 903 B4, Germany.
37. Huber, F., Kalbitzer, H.R. and Kremer, W. (2011) *Process for the determination of lipoproteins in body fluids*. US7927878, USA.
38. Huber, F., Kalbitzer, H.R. and Kremer, W. (2011) *Process for the determination of lipoproteins in body fluids*. AU2005250571, Australia.
39. Altmaier, E., Ramsay, S.L., Graber, A., Mewes, H.W., Weinberger, K.M. and Suhre, K. (2008) Bioinformatics analysis of targeted metabolomics—uncovering old and new tales of diabetic mice under medication. *Endocrinology*, **149**, 3478–3489.
40. Suhre, K., Shin, S.Y., Petersen, A.K., Mohney, R.P., Meredith, D., Wagele, B., Altmaier, E., Deloukas, P., Erdmann, J., Grundberg, E. *et al.* (2011) Human metabolic individuality in biomedical and pharmaceutical research. *Nature*, **477**, 54–60.
41. Paradis, E., Claude, J. and Strimmer, K. (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
42. R Development Core Team. (2010) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
43. Suzuki, R. and Shimodaira, H. (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, **22**, 1540–1542.
44. Shimodaira, H. (2002) An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.*, **51**, 492–508.