



Published in final edited form as:

Circ Cardiovasc Genet. 2013 April ; 6(2): 171–183. doi:10.1161/CIRCGENETICS.112.964619.

Genome-Wide Association Study Identifies Novel Loci Associated With Concentrations of Four Plasma Phospholipid Fatty Acids in the De Novo Lipogenesis Pathway: Results from the CHARGE Consortium

Jason H.Y. Wu, PhD^{1,2}, Rozenn N. Lemaitre, PhD, MPH³, Ani Manichaikul, PhD⁴, Weihua Guan, PhD⁵, Toshiko Tanaka, PhD⁶, Millennia Foy, PhD⁷, Edmond K. Kabagambe, PhD⁸, Luc Djousse, MD, ScD^{9,10}, David Siscovick, MD, MPH^{3,11}, Amanda M. Fretts, PhD, MPH³, Catherine Johnson, PhD, MPH³, Irena B. King, PhD¹², Bruce M. Psaty, MD, PhD^{3,13,14}, Barbara McKnight, PhD¹⁵, Stephen S. Rich, PhD¹⁶, Yii-Der I. Chen, PhD¹⁷, Jennifer A. Nettleton, PhD¹⁸, Weihong Tang, MD, PhD¹⁹, Stefania Bandinelli, MD²⁰, David R. Jacobs Jr, PhD¹⁹, Brian L. Browning, PhD²¹, Cathy C. Laurie, PhD¹⁵, Xiangjun Gu, PhD⁷, Michael Y. Tsai, PhD²², Lyn M. Steffen, PhD, MPH¹⁹, Luigi Ferrucci, MD, PhD⁶, Myriam Fornage, PhD^{7,18}, and Dariush Mozaffarian, MD, DrPH^{1,23}

¹Dept of Epidemiology & Nutrition, Harvard School of Public Health, Boston, MA

²School of Medicine & Pharmacology, University of Western Australia, Perth, Australia

³Cardiovascular Health Research Unit, Dept of Medicine, University of Washington, Seattle, WA

⁴Center for Public Health Genomics, Division of Biostatistics & Epidemiology, University of Virginia, Charlottesville, VA

⁵Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN

⁶Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD

⁷Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX

⁸Dept of Epidemiology, University of Alabama at Birmingham, Birmingham, AL

⁹Division of Aging, Dept of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

¹⁰Boston VA Healthcare System, Boston, MA

¹¹Dept of Epidemiology, School of Public Health, University of Washington, Seattle, WA

¹²Dept of Internal Medicine, University of New Mexico, Albuquerque, NM

¹³Depts of Epidemiology & Health Services, University of Washington, Seattle, WA

¹⁴Group Health Research Institute, Group Health Cooperative, Seattle, WA

¹⁵Dept of Biostatistics, School of Public Health, University of Washington, Seattle, WA

¹⁶Center for Public Health Genomics, University of Virginia, Charlottesville, VA

Corresponding author: Jason Wu, PhD, Harvard School of Public Health, Kresge Building, 913, 677 Huntington Avenue, Boston, MA 02115, Tel: 1-617-432-7981, Fax: 1-617-566-7805, jasonwu@hsph.harvard.edu.

Conflict of Interest Disclosures: None

¹⁷Medical Genetics Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA

¹⁸Division of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas Health Science Center at Houston, Houston, TX

¹⁹Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

²⁰Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy

²¹Dept of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA

²²Laboratory Medicine & Pathology, University of Minnesota, Minneapolis, MN

²³Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Abstract

Background—Palmitic acid(16:0), stearic acid(18:0), palmitoleic acid(16:1n-7), and oleic acid(18:1n-9) are major saturated and mono-unsaturated fatty acids that affect cellular signaling and metabolic pathways. They are synthesized via de novo lipogenesis (DNL) and are the main saturated and mono-unsaturated fatty acids in the diet. Levels of these fatty acids have been linked to diseases including type 2 diabetes and coronary heart disease.

Methods and Results—Genome-wide association studies were conducted in 5 population-based cohorts comprising 8,961 participants of European ancestry to investigate the association of common genetic variation with plasma levels of these four fatty acids. We identified polymorphisms in 7 novel loci associated with circulating levels of one or more of these fatty acids. *ALG14* (asparagine-linked glycosylation 14 homolog) polymorphisms were associated with higher 16:0($P=2.7\times 10^{-11}$) and lower 18:0($P=2.2\times 10^{-18}$). *FADS1* and *FADS2* (desaturases) polymorphisms were associated with higher 16:1n-7($P=6.6\times 10^{-13}$) and 18:1n-9($P=2.2\times 10^{-32}$), and lower 18:0($P=1.3\times 10^{-20}$). *LPGATI* (lysophosphatidylglycerol acyltransferase) polymorphisms were associated with lower 18:0($P=2.8\times 10^{-9}$). *GCKR*(glucokinase regulator, $P=9.8\times 10^{-10}$) and *HIF1AN*(factor inhibiting hypoxia-inducible factor-1, $P=5.7\times 10^{-9}$) polymorphisms were associated with higher 16:1n-7, whereas *PKD2LI*(polycystic kidney disease 2-like 1, $P=5.7\times 10^{-15}$) and a locus on chromosome 2(not near known genes) were associated with lower 16:1n-7($P=4.1\times 10^{-8}$).

Conclusion—Our findings provide novel evidence that common variations in genes with diverse functions, including protein-glycosylation, polyunsaturated fatty acid metabolism, phospholipid modeling, and glucose- and oxygen-sensing pathways, are associated with circulating levels of four fatty acids in the DNL pathway. These results expand our knowledge of genetic factors relevant to DNL and fatty acid biology.

Keywords

epidemiology; fatty acids; genome-wide association study

De novo lipogenesis (DNL) is an endogenous pathway for lipid synthesis, largely occurring in the liver and stimulated by carbohydrate intake and alcohol use ¹. The major products of DNL include the saturated fatty acids, palmitic acid (16:0) and stearic acid (18:0); and the monounsaturated fatty acids, palmitoleic acid (16:1n-7) and oleic acid (18:1n-9) (Figure 1) ². These fatty acids are also consumed in the diet, but endogenous synthesis appears to contribute substantially to their circulating levels, based on short-term metabolic studies that demonstrated increased plasma levels (including in triglycerides and plasma phospholipids) of these fatty acids following high-carbohydrate low-fat diets or acute alcohol intake ³⁻⁵.

Also, in prior work, both higher carbohydrate consumption and higher alcohol use were independently associated with higher levels of these fatty acids in free-living adults⁶. These relationships support the link between DNL and plasma fatty acid levels seen in short-term controlled trials, and indicate that DNL could influence circulating fatty acid levels at usual ranges of dietary exposures.^{6,7} These fatty acids also show weak to moderate associations with direct dietary consumption⁸⁻¹⁰.

In addition to conversion of excess carbohydrate or protein energy into fatty acids for storage as triglycerides, experimental evidence suggests that DNL and its fatty acid products affect multiple signaling and metabolic pathways. In mice, DNL generates endogenous ligands for peroxisome proliferator-activated receptor- α in the liver, a key transcription factor regulating glucose and lipid homeostasis^{11, 12}. Adipose-DNL-derived palmitoleic acid has been shown to downregulate liver fatty acid synthesis and increase skeletal muscle insulin sensitivity in mice¹³. Also, excess hepatic DNL alters hepatic endoplasmic reticulum membrane lipid composition and calcium handling in obese mice, and may represent a mechanism for the development of chronic stress in this important organelle¹⁴. These and other animal studies suggest an influence of DNL and its fatty acid products on crucial physiological functions and pathologic conditions, including insulin sensitivity^{13, 15}, feedback on hepatic fatty acid synthesis¹³, modulation of food intake and energy balance^{16, 17}, and development of hepatic steatosis^{18, 19}. Consistent with their importance in metabolic regulation, circulating levels of fatty acids in the DNL pathway have been linked to risk of several major cardiometabolic diseases, including type 2 diabetes²⁰⁻²³, hypertension²⁴, coronary heart disease (CHD)^{25, 26}, heart failure²⁷, and sudden cardiac arrest^{6, 28}.

Despite these findings suggesting importance of these fatty acids, the metabolic pathways that regulate and determine their circulating concentrations are not understood. Several enzymes involved in their initial synthesis and interconversion are recognized (Figure 1)²⁹; however, potential influences of other enzyme activities or regulatory pathways on circulating levels are unknown. It is likely that genetic variation in other processes, such as intestinal absorption, plasma phospholipid remodeling, dietary preferences, or other unknown pathways, could each affect these fatty acids. Both family and twin studies suggest high heritability of circulating levels of these fatty acids^{30, 31}. The genes that account for the observed familial aggregation of these circulating fatty acid levels remain unknown. We hypothesized that an unbiased investigation of genetic determinants of their circulating levels may identify novel genetic determinants and inform novel metabolic and regulatory pathways and, ultimately, links to disease risk. We therefore carried out a genome-wide association study (GWAS) to identify novel genetic loci contributing to variation in circulating levels of 16:0, 18:0, 16:1n-7, and 18:1n-9.

Methods

Ethics Statement

All cohort participants gave written informed consent, including consent to participate in genetic studies. All studies received approval from local ethical oversight committees.

Study cohorts

2 cohort studies in the Cohort for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium contributed data to the present analysis: the Atherosclerosis Risk in Communities (ARIC) Study and the Cardiovascular Health Study (CHS). Data were also obtained from 3 additional cohorts: the Coronary Artery Risk Development in Young Adults (CARDIA), the Invecchiare in Chianti (InCHIANTI) Study, and the Multi-Ethnic Study of

Atherosclerosis (MESA). Only subjects of European ancestry from each cohort were included in the current analysis.

Fatty Acid Measurements

Details of fatty acid analysis are provided in Supplemental Methods and Supplement Table 1. Briefly, in all cohorts except InCHIANTI, fasting plasma phospholipids were isolated by thin-layer chromatography; fatty acids were subsequently quantified by gas chromatography³². In InCHIANTI, a similar gas chromatography method was used to measure total fasting plasma fatty acids. Levels of 16:0, 16:1n-7, 18:0 and 18:1n-9 were expressed as % of total fatty acids. While several other fatty acids are also in the DNL pathway (e.g. 14:0, 16:1n-9, 18:1n-7), we focused on the 4 key fatty acids (16:0, 16:1n-7, 18:0, and 18:1n-9) that are both major products of DNL^{33,34}, and were available in all the cohorts in this consortium.

Imputation and Cohort-Specific Genome-wide Association Analysis

Details on genotyping and imputation in each cohort are provided in the Supplemental Methods and Supplemental Table 1. Briefly, in each cohort, genotyping was done separately using high-density SNP marker platforms (ARIC, CARDIA and MESA - Affymetrix 6.0, CHS - Illumina 370, InChianti - Illumina 550). Samples with call rates below 95% (ARIC, CARDIA, MESA), or 97% (CHS, InCHIANTI) at genotyped markers were excluded. Genotypes were imputed to approximately 2.5 million HapMap SNPs by using either MACH³⁵ (ARIC, InCHIANTI), BEAGLE³⁶ (CARDIA), BIMBAM³⁷ (CHS), or IMPUTE³⁸ (MESA). SNPs for which testing Hardy Weinberg equilibrium resulted in $p < 10^{-5}$ (CHS, ARIC) or $p < 10^{-4}$ (CARDIA, InCHIANTI) were excluded from imputation. SNPs with minor allele frequency (MAF) $< 1\%$ were excluded from the meta-analyses.

Association analysis between genotype and each fatty acid was done separately within each study cohort according to a pre-specified plan. All studies conducted linear regression analysis using an additive genetic model, *i.e.* regression of phenotype on the number of reference alleles, or equivalently the imputed dosage for imputed genotypes. All analyses were adjusted for age, sex, and site of recruitment where appropriate, and used robust standard errors³⁹. We assessed the SNP-fatty acid GWAS results from each cohort for potential population substructure by examining plots of observed vs expected p-values for each set of GWAS results. We corrected for possible population stratification in 3 of the cohorts (CARDIA, CHS, and MESA) by including the 10 first principal components as covariates in the SNP-fatty acid GWAS. Principal component analysis was not applied in ARIC and InCHIANTI because the analyses in these cohorts show inflation factors very close to 1, however, genomic control correction was applied to each study prior to the meta-analysis to additionally minimize potential confounding by population stratification. Inflation factors (λ) ranged from 0.993-1.042 (16:0), 0.993-1.035 (18:0), 0.995-1.085 (16:1n-7), and 1.001-1.073 (18:1n-9), suggesting minimal population stratification after the adjustments.

Meta-Analysis

For each SNP and fatty acid, study-specific genome-wide association results (treating the minor allele of each SNP as the risk allele) were combined using inverse-variance weighted meta-analysis in METAL (www.sph.umich.edu/csg/abecasis/metal.htm). P-values less than 5×10^{-8} were considered significant. Because total plasma levels of the fatty acids of interest in InChianti differed substantially from their plasma phospholipid levels in the other cohorts (Table 1), we performed a z-score based meta-analysis of each fatty acid with the 5 cohorts, as a sensitivity analysis. The proportion of fatty acid variance explained by a particular variant allele was calculated for each cohort from the formula $\text{corr}(Y, \hat{Y})^2 \cong (\beta^2 * 2 * \text{MAF}(1 -$

MAF)/Var(Y), where β is the regression coefficient for one copy of the allele, MAF is the minor allele frequency and Var(Y) is the variance of the fatty acid (Online Supplemental Material).

Interaction analyses

We tested 20 interactions using cross products in the linear regression models. We investigated interactions of the 10 most associated SNPs with a) dietary carbohydrate intake (as substitution for an equal amount of % of total energy from fat, entered as a continuous linear variable) and b) habitual alcohol intake (grams/day + 1, log transformed, and entered as a continuous linear variable) on the outcomes of 16:0, 16:1n-7, 18:0 and 18:1n-9. Interaction coefficients from cohort-specific analyses were meta-analyzed. P-values less than 0.0025 (0.05/20 tests) were considered significant for the interactions.

Results

Cohort characteristics

The study sample included a total of 8,961 subjects of European ancestry in the ARIC, CARDIA, CHS, InChianti and MESA cohorts. Table 1 shows sample sizes, subject demographics, selected dietary habits, and circulating fatty acid levels in the 5 cohorts. Participants comprised mostly middle-aged to older individuals (mean age across the cohorts ranged from mid-40s to mid-70s), and approximately 50% were men. Typical for middle-aged and older US populations, subjects obtained ~50% of their energy from carbohydrate and reported low to moderate alcohol consumption^{40, 41}. As percent of total fatty acids, mean levels of 16:0 varied between the cohorts from 22.5% to 25.8%; 16:1n-7, from 0.51% to 2.4%; 18:0, from 6.5% to 13.3%; and 18:1n-9, from 7.6% to 25.9%. The InCHIANTI study only measured fatty acids in total plasma, accounting for the larger differences in levels when compared to other cohorts that measured plasma phospholipid fatty acids⁴².

Genome-wide associations of fatty acids in the DNL pathway

The meta-analysis of the genome-wide association results is shown in Figure 2. SNPs in multiple novel genetic loci were associated with these fatty acids. For most of these loci, many highly correlated SNPs reached genome-wide significance ($P < 5 \times 10^{-8}$, 379 total SNPs; Table 2 and Supplementary Tables 2-5). For the results below, directionality is reported for minor alleles

Variants alleles of SNPs on a chromosome 1 locus that contained the *ALG14*, *RWDD3*, *CNN3*, and *TMEM56* genes were associated with higher 16:0 (most associated SNP: rs2391388, $P = 2.7 \times 10^{-11}$, Table 2 and Figure 3) and lower 18:0 (most associated SNP: rs6675668, $P = 2.2 \times 10^{-18}$, Table 2 and Figure 4A). These two most significant SNPs were both in the *ALG14* gene and in strong linkage disequilibrium ($r^2 = 0.875$).

Variant alleles of SNPs at two additional loci were associated with lower levels of 18:0 (Table 2 and Figure 4). The first locus was on chromosome 1 that contained *LPGAT1*, a gene involved in phospholipid metabolism (most associated SNP: rs11119805, $P = 2.8 \times 10^{-9}$, Figure 4B) and another on chromosome 11 that contained the *C11orf9/10*, *FEN1*, and desaturase genes (*FADS1* and *FADS2*) (most associated SNP: rs102275, $P = 1.3 \times 10^{-20}$, Figure 4C). Variant alleles of SNPs at the *FADS1/2* gene cluster was also associated with higher 18:1n-9 (Figure 5) and 16:1n-7 (Figure 6A); with rs102275 also being the most significantly associated SNP ($P = 2.2 \times 10^{-32}$ and 6.6×10^{-13} for 18:1n-9 and 16:1n-7, respectively). For 18:1n-9, the SNPs that reached genome-wide significance extended to *FADS3* (another putative desaturase gene) (Figure 5 and Supplemental Table S1).

Four additional loci were associated with circulating levels of 16:1n-7. One locus was on chromosome 10 where a single SNP (rs603424, $P = 5.7 \times 10^{-15}$, Table 2 and Figure 6B) in the gene *PKD2L1* was associated with lower 16:1n-7. Variant alleles in the other 3 loci were all associated with higher 16:1n-7. One locus was on chromosome 10 (most associated SNP: rs11190604, $P = 5.7 \times 10^{-9}$, Figure 6C), and contained the *HIF1AN*, *NDUFB8*, *SEC31B*, and *WNT8B* genes. Two separate regions were also found on chromosome 2: one contained *GCKR* (most associated SNP: rs780093, $P = 9.8 \times 10^{-10}$, Table 2 and Figure 6D), a gene with key regulatory roles in liver glucose and lipid metabolism. The second locus (most associated SNP: rs6722456, $P = 4.1 \times 10^{-8}$) had no known gene in the region.

The genome-wide significant SNPs were either genotyped or had good to excellent imputation quality (Supplemental Table 6). Across the 5 individual cohorts, the directions of association between the top SNPs at each locus and the fatty acid concentrations were consistent (Figure 7, study specific beta-coefficients and standard errors are available on request). The magnitude and direction of associations were generally consistent across all 5 cohorts. In cases where moderate heterogeneity were present, this was largely due to different findings in the InCHIANTI cohort. This heterogeneity was due to differences in the magnitudes of the gene-fatty acid associations in InCHIANTI versus the other cohorts, rather than differing directions of associations.

Sensitivity analysis

Using inverse-variance-weighted meta-analysis assumed that the fatty acids were measured on the same scale in all the cohorts. However, the InCHIANTI cohort measured total plasma fatty acids while the other 4 cohorts measured plasma phospholipid fatty acids, and whether the scale of these different measurements was the same was uncertain. To assess whether the findings depended on the meta-analytical technique, we performed z-score based meta-analysis. Results were not altered appreciably for all of the most associated SNPs in Table 2 (not shown) with the exception of rs6722456 (z-score based meta-analysis $P = 3.7 \times 10^{-5}$). Furthermore, exclusion of InCHIANTI from each meta-analysis did not materially alter the top SNP-fatty acid associations (results not shown).

We also found little evidence for interaction between dietary carbohydrate intake or alcohol use and any of the identified SNPs in relation to levels of 16:0, 16:1n-7, 18:0, or 18:1n-9 (P interaction > 0.0025 each, the Bonferroni-adjusted significance threshold).

Discussion

We report the first evidence of genome-wide significant associations of 7 novel loci associated with one or more of four plasma phospholipid fatty acids in the DNL pathway. These results are based upon meta-analysis of GWAS of circulating fatty acids in nearly 9,000 adults of European ancestry. The directions of association of the most associated SNPs are shown in Figure 8. Two of the identified loci have opposite associations with adjacent fatty acids in the metabolic pathway, *i.e.*, *ALG14* with higher 16:0 and lower 18:0, and *FADS1/2* with lower 18:0 and higher 18:1n-9. This suggests the genes have potentially distinct effects on metabolic conversion of these paired fatty acids. A third locus, *LPGATI*, was also associated with levels of 18:0, and four other loci (*GCKR*, *HIF1AN*, *PKD2L1*, one on chromosome 2 not near known genes) as well as *FADS1/2* were associated with variation in levels of 16:1n-7.

Plasma phospholipid fatty acid levels are complex physiological traits and likely influenced by a variety of exogenous and endogenous factors. The biologic pathways represented by the novel loci identified in this study, and their potential relations to fatty acid metabolism and DNL, are intriguing. For example, we found minor alleles in *FADS1/2* on chromosome

11 to be associated with both lower 18:0 and higher 18:1n-9, as well as higher 16:1n-7. Our findings confirm a prior candidate gene study that observed associations of *FADS1/2* variation and erythrocyte phospholipid 16:1n-7 levels among 4457 pregnant women⁴³, and further extend these findings by demonstrating associations with 18:0 and 18:1n-9, for which statistical power may have been previously limited. The opposite directions of association of these SNP alleles with 18:0 and 18:1n-9 suggest a potential effect on the corresponding Δ -9 desaturation reaction, which is predominantly catalyzed by stearoyl-CoA desaturase (SCD-1)². However, *FADS1/2* encodes the Δ -5 and Δ -6 fatty acid desaturases, involved in polyunsaturated fatty acids biosynthesis⁴⁴, and not the Δ -9 desaturase. We recently reported that genetic variation in *FADS1/2*, including SNPs in high LD ($r^2 = 0.9$) with the currently identified top SNP in this locus, is associated with circulating phospholipid n-3 fatty acids, including higher alpha-linolenic acid and lower eicosapentaenoic acid⁶. Further work is needed to investigate how genetic variation in *FADS1/2* could affect circulating concentrations of 16:1n-7, 18:0, and 18:1n-9. Finally, several SNPs in *FADS3* was associated with 18:1n-9 at genome-wide significance (Supplemental Table 1). Given the presence of multiple LD blocks in the *FADS1/2/3* region, the associations could also be due to *FADS3*, a putative desaturase gene whose function is currently unknown. Our results suggest future studies should investigate if *FADS3* possesses Δ -9 desaturase activities.

Our findings concerning *ALG14* are consistent with a prior GWAS study in ~4000 European subjects that found an association of *ALG14* variation with lysophosphatidylcholine (a sub-class of phospholipids) 18:0⁴⁵, and extend these findings by showing a relationship to total phospholipid 18:0 and 16:0. *ALG14* encodes the asparagine (N)-linked glycosylation 14 homolog, a subunit of the UDP-N-acetyl glucosamine transferase (UDP- N-GlcNAc) protein. In the endoplasmic reticulum, UDP- N-GlcNAc contributes to biosynthesis of lipid-linked oligosaccharides that are essential substrates for N-linked glycosylation of proteins^{46, 47}. The endoplasmic reticulum is the predominant site of the elongation of long-chain fatty acids⁴⁸. Furthermore, inhibition of N-linked glycosylation in cultured rat hepatocytes enhances activation of sterol regulatory element-binding protein-1 (SREBP-1, a transcription factor regulating the expression of lipogenic enzymes), and increases downstream expression of DNL enzymes including fatty acid synthase (FAS, which carries out the initial steps in DNL to produce 16:0 as the major product) and stearoyl CoA desaturase (SCD-1), which catalyzes Δ -9 desaturation of 16:0 to 16:1n-7 and of 18:0 to 18:1n-9)¹⁸. When viewed in the context of these prior studies, our findings of an association between genetic variation in *ALG14* and circulating 16:0 and 18:0 levels highlights the need to investigate whether *ALG14* affects fatty acid metabolism and/or DNL, for example through changes in N-glycosylation of proteins synthesized in the endoplasmic reticulum.

Variation in *LPGAT1*, which encodes for a lysophosphatidylglycerol acyltransferase⁴⁹, was associated with lower 18:0. After their initial synthesis, phospholipids such as phosphatidylglycerol undergo rapid deacylation-reacylation modeling (Lands' cycle) that alters the fatty acid present at their sn-2 position, which may influence phospholipid properties⁵⁰. Recent molecular studies suggest that the *LPGAT1* protein preferentially transfers 16:0, 18:0, and 18:1n-9 to the sn-2 position of lysophosphatidylglycerol to generate phosphatidylglycerol, and thus contribute to the reacylation pathway⁴⁹. The novel genetic association observed in this study, as well as the growing evidence of the potentially important role of phosphatidylglycerol and phospholipid remodeling in cell signaling and metabolism⁵¹⁻⁵³, should encourage further research into genetic determinants of phosphatidylglycerol fatty acid composition and function.

We found 5 distinct loci associated with 16:1n-7, far more than for any of the other fatty acids evaluated. Because relatively few dietary sources of 16:1n-7 exist⁵⁴, metabolic factors may be particularly important in determining its circulating concentrations. Among the identified SNPs on chromosomes 2 and 10, *GCKR* and *HIF1AN* stand out as potential genes of interest. In *GCKR*, the T allele of rs780093 (the most significantly associated SNP with 16:1n-7) is in strong LD with the T allele of rs1260326 ($r^2=0.9$), a non-synonymous variant (P446L) that has been consistently associated with elevated triglycerides and reduced fasting plasma glucose^{55, 56}. *GCKR* encodes glucokinase regulatory protein, which inhibits the activity of glucokinase, a key protein with many roles including glucose sensing in liver and pancreas⁵⁷. Experimentally, P446L-*GCKR* has reduced inhibitory activity towards glucokinase, which would be predicted to increase glycolytic flux and production of malonyl-CoA, a key substrate for DNL pathway and triglyceride synthesis⁵⁶. Such genetically influenced increased availability of substrate for DNL might account for the observed association with 16:1n-7, a hypothesis that warrants testing in future investigations.

HIF1AN encodes a protein which modulates the activity of HIF-1 α , a transcription factor regulating cellular metabolic responses to oxygen availability⁵⁸. In mice, partial deficiency of HIF-1 α attenuates the induction by intermittent hypoxia of fatty acid and triglyceride biosynthesis, lowering SREBP-1 and SCD-1 activity and increasing the serum ratio of 16:1n-7 to 16:0^{59, 60}. Knockout of the *HIF1AN* gene in mice resulted in decreased expression of lipogenic genes (e.g. *Scd-1*) in response to a high-fat, high-calorie diet⁶¹. Our present results support a link between *GCKR* and *HIF1AN* with DNL in humans.

PKD2L1, which encodes for polycystic kidney disease 2-like 1 protein⁶², was also associated with lower 16:1n-7 levels. In support of our finding, a prior GWAS analysis in European populations reported similar associations of *PKD2L1* with lysophosphatidylcholine (a sub-class of phospholipids) 16:1 n-7⁴⁵. This protein has been suggested to function as a calcium-regulated non-selective cation channel with a role in detecting sour taste in humans^{63, 64}. We are unaware of experimental studies suggesting its involvement in pathways related to 16:1n-7 or DNL. 16:1n-7 levels were also associated with several other SNPs in chromosome 10 genes, including *NDUFB8* that encodes NADH dehydrogenase 1 beta sub-complex, *SEC31B* that encodes SEC31 homolog B, and *WNT8B* that encodes wingless-type MMTV integration site family member 8B. We could not identify prior studies suggesting a role of these genes in DNL or fatty acid metabolism. Further investigation of these potential effects may elucidate novel regulatory pathways and functions.

Finally, we found a locus on chromosome 2 (top SNP: rs6722456) associated with 16:1n-7 that was not in close vicinity of known genes. The direction of association was consistent in 4 of 5 cohorts, but nevertheless the relevance of this finding is as yet unknown. A strength of this current analysis is that all of the participants in this consortium were from random subsets of cohorts, rather than nested case-controls studies, thus reducing the chance of selection bias. An additional strength of genetic studies such as this one is the absence of potential confounding from environmental or lifestyle factors. Our results were consistent across the cohorts, with minor heterogeneity present for InCHIANTI. We note that all other cohorts assessed fatty acids in plasma phospholipids, whereas fatty acids were measured in total plasma in InCHIANTI; this difference could, at least in part, account for some of this heterogeneity. We additionally confirmed the consistency of our results as exclusion of InCHIANTI from the meta-analysis had minimal to no effect on the observed associations. In our recent GWAS of plasma phospholipid n-3 fatty acids, the majority of identified variants were in genes encoding known regulatory enzymes for these fatty acids³². In contrast, in the present study, variation in genes encoding proteins in the DNL pathway (e.g,

stearoyl CoA desaturase) was not associated with DNL fatty acids. These results suggest that metabolic pathways other than DNL may influence circulating levels of these fatty acids, and also highlight the unique strength of appropriately powered, non-hypothesis-driven GWAS for identifying novel genetic polymorphisms that affect circulating fatty acids. Such approaches can be complemented by pathway-driven analyses in future work, e.g., focusing on known fatty acid regulatory genes that may not have been identified herein due to small effects or very low frequency. Potential limitations should be considered. The effect of polymorphisms on protein function is unknown and the identified variants explained little of the variance in fatty acid levels. Thus, our results highlight the need for additional studies to identify other factors, including direct dietary consumption, gene-environment interactions, and other endogenous metabolic processes that influence these fatty acids, and possible heterogeneity across populations. However, as with most other GWAS, we believe that the findings of our analysis are most relevant to identify new biologic pathways that influence these fatty acids and possibly DNL, rather than to account for population variation⁶⁵. It should be noted that each identified genetic loci contained many SNPs in high LD and we do not know which are the causal variants. Additionally, while the top findings were highly statistically significant, for SNPs with *P*-values close to the genome-wide threshold should be interpreted with caution, and future replication efforts are required to confirm these associations. Another limitation was that we focused on fatty acids which were major products of DNL and that were available in all cohorts in the consortium. Several other fatty acids in the DNL pathway (e.g. myristic acid) were therefore not investigated and should be the focus of future studies. Finally, each of the cohorts measured different number of fatty acids, which potentially introduced artificial variation in % fatty acid abundance in cross-cohort comparisons. However, the magnitude and direction of associations were mostly consistent across all 5 cohorts, suggesting the influence of different number of fatty acids measured had minimal effects on the SNP-fatty acid associations.

In conclusion, in the first reported GWAS of circulating levels of these fatty acids, we identified multiple novel genetic variants in genes with functions in protein *N*-glycosylation, polyunsaturated fatty acid metabolism, phospholipid modeling, and glucose and oxygen sensing pathways. Given the multi-faceted cellular functions of these fatty acids and of DNL, as well as their potential influence on cardiometabolic diseases, these findings inform new directions for investigation of how these genes and related molecular mechanisms may alter fatty acid synthesis and metabolism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the other investigators, the staff and the participants of the ARIC study, the CARDIA study, the CHS study, the MESA study, and the InCHIANTI study for their important contributions. A full list of principal CHS investigators and institutions can be found at <http://www.chs-nhlbi.org/pi.htm>. A full list of principal CARDIA investigators and institutions can be found at http://www.cardia.dopm.uab.edu/o_pain.htm. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>. The authors acknowledge the essential role of the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Consortium in development and support of this manuscript. CHARGE members include NHLBI's Atherosclerosis Risk in Communities (ARIC) Study, NIA's Iceland Age, Gene/Environment Susceptibility (AGES) Study, NHLBI's Cardiovascular Health Study (CHS) and Framingham Heart Study (FHS), and the Netherland's Rotterdam Study (RS).

Funding Sources: The *Atherosclerosis Risk in Communities Study* is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C,

HHSN268201100011C, HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The *Cardiovascular Health Study* research reported in this article was supported by the National Heart, Lung, and Blood Institute contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, and grants HL080295, HL075366, HL087652, HL105756 with additional contribution from National Institute of Neurological Disorders and Stroke. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the National Institute on Aging. See also <http://www.chs-nhlbi.org/pi.htm>. DNA handling and genotyping was supported in part by National Center for Research Resources CTSI grant UL1RR033176 and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The *Coronary Artery Risk Development in Young Adults (CARDIA)* study is funded by contracts N01-HC-95095, N01-HC-48047, N01-HC-48048, N01-HC-48049, N01-HC-48050, N01-HC-45134, N01-HC-05187, N01-HC-45205, and N01-HC-45204 from the National Heart, Lung and Blood Institute to the CARDIA investigators. Genotyping of the CARDIA participants was supported by grants U01-HG-004729, U01-HG-004446 and U01-HG-004424 from the National Human Genome Research Institute and R01-HL-084099 from the National Heart, Lung and Blood Institute to MF. Statistical analyses and fatty acid measures were funded by R01-HL-084099 from the National Heart, Lung and Blood Institute to MF. The *InCHIANTI Study* baseline (1998-2000) was supported as a “targeted project” (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland. The *Multi-Ethnic Study of Atherosclerosis (MESA)* and MESA SHARE were supported by contracts N01-HC-95159 through N01-HC-95169 and RR-024156 from the National Heart, Lung and Blood Institute. Funding for MESA SHARE genotyping was provided by NHLBI Contract N02-HL-6-4278. MESA Family is conducted and supported in collaboration with MESA investigators; support is provided by grants and contracts R01HL071051, R01HL071205, R01HL071250, R01HL071251, R01HL071252, R01HL071258, R01HL071259. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>. Dr. Jennifer Nettleton is supported by a K01 from the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (5K01DK082729-02).

References

- Hellerstein MK, Schwarz JM, Neese RA. Regulation of hepatic de novo lipogenesis in humans. *Annu Rev Nutr.* 1996; 16:523–557. [PubMed: 8839937]
- Paton CM, Ntambi JM. Biochemical and physiological function of stearoyl-coa desaturase. *Am J Physiol Endocrinol Metab.* 2009; 297:E28–37. [PubMed: 19066317]
- King IB, Lemaitre RN, Kestin M. Effect of a low-fat diet on fatty acid composition in red cells, plasma phospholipids, and cholesterol esters: Investigation of a biomarker of total fat intake. *Am J Clin Nutr.* 2006; 83:227–236. [PubMed: 16469979]
- Raatz SK, Bibus D, Thomas W, Kris-Etherton P. Total fat intake modifies plasma fatty acid composition in humans. *J Nutr.* 2001; 131:231–234. [PubMed: 11160538]
- Siler SQ, Neese RA, Hellerstein MK. De novo lipogenesis, lipid kinetics, and whole-body lipid balances in humans after acute alcohol consumption. *Am J Clin Nutr.* 1999; 70:928–936. [PubMed: 10539756]
- Wu JH, Lemaitre RN, Imamura F, King IB, Song X, Spiegelman D, et al. Fatty acids in the de novo lipogenesis pathway and risk of coronary heart disease: The cardiovascular health study. *Am J Clin Nutr.* 2011; 94:431–438. [PubMed: 21697077]
- Simon JA, Fong J, Bernert JT Jr, Browner WS. Relation of smoking and alcohol consumption to serum fatty acids. *Am J Epidemiol.* 1996; 144:325–334. [PubMed: 8712189]
- Hodge AM, Simpson JA, Gibson RA, Sinclair AJ, Makrides M, O’Dea K, et al. Plasma phospholipid fatty acid composition as a biomarker of habitual dietary fat intake in an ethnically diverse cohort. *Nutr Metab Cardiovasc Dis.* 2007; 17:415–426. [PubMed: 16962297]
- Saadatian-Elahi M, Slimani N, Chajes V, Jenab M, Goudable J, Biessy C, et al. Plasma phospholipid fatty acid profiles and their association with food intakes: Results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr.* 2009; 89:331–346. [PubMed: 19056549]

10. Sun Q, Ma J, Campos H, Hankinson SE, Hu FB. Comparison between plasma and erythrocyte fatty acid content as biomarkers of fatty acid intake in us women. *Am J Clin Nutr.* 2007; 86:74–81. [PubMed: 17616765]
11. Chakravarthy MV, Pan Z, Zhu Y, Tordjman K, Schneider JG, Coleman T, et al. “New” hepatic fat activates pparalpha to maintain glucose, lipid, and cholesterol homeostasis. *Cell Metab.* 2005; 1:309–322. [PubMed: 16054078]
12. Chakravarthy MV, Lodhi IJ, Yin L, Malapaka RR, Xu HE, Turk J, et al. Identification of a physiologically relevant endogenous ligand for pparalpha in liver. *Cell.* 2009; 138:476–488. [PubMed: 19646743]
13. Cao H, Gerhold K, Mayers JR, Wiest MM, Watkins SM, Hotamisligil GS. Identification of a lipokine, a lipid hormone linking adipose tissue to systemic metabolism. *Cell.* 2008; 134:933–944. [PubMed: 18805087]
14. Fu S, Yang L, Li P, Hofmann O, Dicker L, Hide W, et al. Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature.* 2011; 473:528–531. [PubMed: 21532591]
15. Lodhi IJ, Wei X, Semenkovich CF. Lipoexpediency: De novo lipogenesis as a metabolic signal transmitter. *Trends Endocrinol Metab.* 2011; 22:1–8. [PubMed: 20889351]
16. Proulx K, Cota D, Woods SC, Seeley RJ. Fatty acid synthase inhibitors modulate energy balance via mammalian target of rapamycin complex 1 signaling in the central nervous system. *Diabetes.* 2008; 57:3231–3238. [PubMed: 18776140]
17. Loftus TM, Jaworsky DE, Frehywot GL, Townsend CA, Ronnett GV, Lane MD, et al. Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. *Science.* 2000; 288:2379–2381. [PubMed: 10875926]
18. Kammoun HL, Chabanon H, Hainault I, Luquet S, Magnan C, Koike T, et al. Grp78 expression inhibits insulin and er stress-induced srebp-1c activation and reduces hepatic steatosis in mice. *J Clin Invest.* 2009; 119:1201–1215. [PubMed: 19363290]
19. Basseri S, Austin RC. Er stress and lipogenesis: A slippery slope toward hepatic steatosis. *Dev Cell.* 2008; 15:795–796. [PubMed: 19081067]
20. Wang L, Folsom AR, Zheng ZJ, Pankow JS, Eckfeldt JH. Plasma fatty acid composition and incidence of diabetes in middle-aged adults: The atherosclerosis risk in communities (aric) study. *Am J Clin Nutr.* 2003; 78:91–98. [PubMed: 12816776]
21. Hodge AM, English DR, O’Dea K, Sinclair AJ, Makrides M, Gibson RA, et al. Plasma phospholipid and dietary fatty acids as predictors of type 2 diabetes: Interpreting the role of linoleic acid. *Am J Clin Nutr.* 2007; 86:189–197. [PubMed: 17616780]
22. Kroger J, Zietemann V, Enzenbach C, Weikert C, Jansen EH, Doring F, et al. Erythrocyte membrane phospholipid fatty acids, desaturase activity, and dietary fatty acids in relation to risk of type 2 diabetes in the european prospective investigation into cancer and nutrition (epic)-potsdam study. *Am J Clin Nutr.* 2011; 93:127–142. [PubMed: 20980488]
23. Patel PS, Sharp SJ, Jansen E, Luben RN, Khaw KT, Wareham NJ, et al. Fatty acids measured in plasma and erythrocyte-membrane phospholipids and derived by food-frequency questionnaire and the risk of new-onset type 2 diabetes: A pilot study in the european prospective investigation into cancer and nutrition (epic)-norfolk cohort. *Am J Clin Nutr.* 2010; 92:1214–1222. [PubMed: 20861175]
24. Zheng ZJ, Folsom AR, Ma J, Arnett DK, McGovern PG, Eckfeldt JH. Plasma fatty acid composition and 6-year incidence of hypertension in middle-aged adults: The atherosclerosis risk in communities (aric) study. *Am J Epidemiol.* 1999; 150:492–500. [PubMed: 10472949]
25. Simon JA, Hodgkins ML, Browner WS, Neuhaus JM, Bernert JT Jr, Hulley SB. Serum fatty acids and the risk of coronary heart disease. *Am J Epidemiol.* 1995; 142:469–476. [PubMed: 7677125]
26. Warensjo E, Sundstrom J, Vessby B, Cederholm T, Riserus U. Markers of dietary fat quality and fatty acid desaturation as predictors of total and cardiovascular mortality: A population-based prospective study. *Am J Clin Nutr.* 2008; 88:203–209. [PubMed: 18614742]
27. Yamagishi K, Nettleton JA, Folsom AR. Plasma fatty acid composition and incident heart failure in middle-aged adults: The atherosclerosis risk in communities (aric) study. *Am Heart J.* 2008; 156:965–974. [PubMed: 19061714]

28. Lemaitre RN, King IB, Sotoodehnia N, Knopp RH, Mozaffarian D, McKnight B, et al. Endogenous red blood cell membrane fatty acids and sudden cardiac arrest. *Metabolism*. 2010; 59:1029–1034. [PubMed: 20045147]
29. Green CD, Ozguden-Akkoc CG, Wang Y, Jump DB, Olson LK. Role of fatty acid elongases in determination of de novo synthesized monounsaturated fatty acid species. *J Lipid Res*. 2010; 51:1871–1877. [PubMed: 20228221]
30. Kunesova M, Hainer V, Tvrzicka E, Phinney SD, Stich V, Parizkova J, et al. Assessment of dietary and genetic factors influencing serum and adipose fatty acid composition in obese female identical twins. *Lipids*. 2002; 37:27–32. [PubMed: 11876260]
31. Lemaitre RN, Siscovick DS, Berry EM, Kark JD, Friedlander Y. Familial aggregation of red blood cell membrane fatty acid composition: The kibbutzim family study. *Metabolism*. 2008; 57:662–668. [PubMed: 18442630]
32. Lemaitre RN, Tanaka T, Tang W, Manichaikul A, Foy M, Kabagambe EK, et al. Genetic loci associated with plasma phospholipid n-3 fatty acids: A meta-analysis of genome-wide association studies from the charge consortium. *PLoS Genet*. 2011; 7:e1002193. [PubMed: 21829377]
33. Aarsland A, Wolfe RR. Hepatic secretion of vldl fatty acids during stimulated lipogenesis in men. *J Lipid Res*. 1998; 39:1280–1286. [PubMed: 9643360]
34. Collins JM, Neville MJ, Pinnick KE, Hodson L, Ruyter B, van Dijk TH, et al. De novo lipogenesis in the differentiating human adipocyte can provide all fatty acids necessary for maturation. *J Lipid Res*. 2011; 52:1683–1692. [PubMed: 21677304]
35. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. Mach: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol*. 2010; 34:816–834. [PubMed: 21058334]
36. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet*. 2009; 84:210–223. [PubMed: 19200528]
37. Servin B, Stephens M. Imputation-based analysis of association studies: Candidate regions and quantitative traits. *PLoS Genet*. 2007; 3:e114. [PubMed: 17676998]
38. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007; 39:906–913. [PubMed: 17572673]
39. White H. Maximum likelihood estimation of misspecified models. *Econometrica*. 1982; 50:1–25.
40. Wright, jd; wang, cy. Nchs data brief. Hyattsville, md; National center for health statistics: 2010. Trends in intake of energy and macronutrients in adults from 1999-2000 through 2007-2008.
41. Hvidtfeldt UA, Tolstrup JS, Jakobsen MU, Heitmann BL, Gronbaek M, O'Reilly E, et al. Alcohol intake and risk of coronary heart disease in younger, middle-aged, and older adults. *Circulation*. 2010; 121:1589–1597. [PubMed: 20351238]
42. Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. *Prog Lipid Res*. 2008; 47:348–380. [PubMed: 18435934]
43. Koletzko B, Lattka E, Zeilinger S, Illig T, Steer C. Genetic variants of the fatty acid desaturase gene cluster predict amounts of red blood cell docosahexaenoic and other polyunsaturated fatty acids in pregnant women: Findings from the avon longitudinal study of parents and children. *Am J Clin Nutr*. 2011; 93:211–219. [PubMed: 21106917]
44. Lattka E, Illig T, Koletzko B, Heinrich J. Genetic variants of the fads1 fads2 gene cluster as related to essential fatty acid metabolism. *Curr Opin Lipidol*. 2010; 21:64–69. [PubMed: 19809313]
45. Demirkan A, van Duijn CM, Ugocsai P, Isaacs A, Pramstaller PP, Liebisch G, et al. Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet*. 2012; 8:e1002490. [PubMed: 22359512]
46. Averbeck N, Keppler-Ross S, Dean N. Membrane topology of the alg14 endoplasmic reticulum udp-glcnae transferase subunit. *J Biol Chem*. 2007; 282:29081–29088. [PubMed: 17686769]
47. Gao XD, Tachikawa H, Sato T, Jigami Y, Dean N. Alg14 recruits alg13 to the cytoplasmic face of the endoplasmic reticulum to form a novel bipartite udp-n-acetylglucosamine transferase required

- for the second step of n-linked glycosylation. *J Biol Chem.* 2005; 280:36254–36262. [PubMed: 16100110]
48. Jakobsson A, Westerberg R, Jacobsson A. Fatty acid elongases in mammals: Their regulation and roles in metabolism. *Prog Lipid Res.* 2006; 45:237–249. [PubMed: 16564093]
49. Yang Y, Cao J, Shi Y. Identification and characterization of a gene encoding human *lpgat1*, an endoplasmic reticulum-associated lysophosphatidylglycerol acyltransferase. *J Biol Chem.* 2004; 279:55866–55874. [PubMed: 15485873]
50. Shindou H, Shimizu T. Acyl-coa:Lysophospholipid acyltransferases. *J Biol Chem.* 2009; 284:1–5. [PubMed: 18718904]
51. Vreken P, Valianpour F, Nijtmans LG, Grivell LA, Plecko B, Wanders RJ, et al. Defective remodeling of cardiolipin and phosphatidylglycerol in barth syndrome. *Biochem Biophys Res Commun.* 2000; 279:378–382. [PubMed: 11118295]
52. Murray NR, Fields AP. Phosphatidylglycerol is a physiologic activator of nuclear protein kinase c. *J Biol Chem.* 1998; 273:11514–11520. [PubMed: 9565565]
53. Schmidt JA, Brown WJ. Lysophosphatidic acid acyltransferase 3 regulates golgi complex structure and function. *J Cell Biol.* 2009; 186:211–218. [PubMed: 19635840]
54. Mozaffarian D, Cao H, King IB, Lemaitre RN, Song X, Siscovick DS, et al. Circulating palmitoleic acid and risk of metabolic abnormalities and new-onset diabetes. *Am J Clin Nutr.* 2010; 92:1350–1358. [PubMed: 20943795]
55. Orho-Melander M, Melander O, Guiducci C, Perez-Martinez P, Corella D, Roos C, et al. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and c-reactive protein but lower fasting glucose concentrations. *Diabetes.* 2008; 57:3112–3121. [PubMed: 18678614]
56. Beer NL, Tribble ND, McCulloch LJ, Roos C, Johnson PR, Orho-Melander M, et al. The p4461 variant in *gckr* associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet.* 2009; 18:4081–4088. [PubMed: 19643913]
57. Iynedjian PB. Molecular physiology of mammalian glucokinase. *Cell Mol Life Sci.* 2009; 66:27–42. [PubMed: 18726182]
58. Weidemann A, Johnson RS. Biology of hif-1alpha. *Cell Death Differ.* 2008; 15:621–627. [PubMed: 18259201]
59. Li J, Bosch-Marce M, Nanayakkara A, Savransky V, Fried SK, Semenza GL, et al. Altered metabolic responses to intermittent hypoxia in mice with partial deficiency of hypoxia-inducible factor-1alpha. *Physiol Genomics.* 2006; 25:450–457. [PubMed: 16507783]
60. Li J, Thorne LN, Punjabi NM, Sun CK, Schwartz AR, Smith PL, et al. Intermittent hypoxia induces hyperlipidemia in lean mice. *Circ Res.* 2005; 97:698–706. [PubMed: 16123334]
61. Zhang N, Fu Z, Linke S, Chicher J, Gorman JJ, Visk D, et al. The asparaginyl hydroxylase factor inhibiting hif-1alpha is an essential regulator of metabolism. *Cell Metab.* 2010; 11:364–378. [PubMed: 20399150]
62. Wu G, Hayashi T, Park JH, Dixit M, Reynolds DM, Li L, et al. Identification of *pkd2l1*, a human *pkd2*-related gene: Tissue-specific expression and mapping to chromosome 10q25. *Genomics.* 1998; 54:564–568. [PubMed: 9878261]
63. Huque T, Cowart BJ, Dankulich-Nagrudny L, Pribitkin EA, Bayley DL, Spielman AI, et al. Sour ageusia in two individuals implicates ion channels of the *ASIC* and *PKD* families in human sour taste perception at the anterior tongue. *PLoS One.* 2009; 4:e7347. [PubMed: 19812697]
64. Kawaguchi H, Yamanaka A, Uchida K, Shibasaki K, Sokabe T, Maruyama Y, et al. Activation of polycystic kidney disease-2-like 1 (*pkd2l1*)-*pkd113* complex by acid in mouse taste cells. *J Biol Chem.* 2010; 285:17277–17281. [PubMed: 20406802]
65. Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. *N Engl J Med.* 2009; 360:1699–1701. [PubMed: 19369661]

Palmitic acid, stearic acid, palmitoleic acid, and oleic acid are major saturated and monounsaturated fatty acids that affect cellular signaling and metabolic pathways. They are synthesized via de novo lipogenesis (DNL) and are the main saturated and monounsaturated fatty acids in the diet. Levels of these fatty acids have been linked to diseases including type 2 diabetes and coronary heart disease. Despite these findings suggesting importance of these fatty acids, the metabolic pathways and genetic variants that regulate and determine their circulating concentrations are not understood. Genome-wide association studies were conducted in 5 population-based cohorts comprising 8,961 participants of European ancestry to investigate the association of common genetic variation with plasma levels of these four fatty acids. We identified polymorphisms in 7 novel loci associated with circulating levels of one or more of these fatty acids. Polymorphisms associated with fatty acid levels were identified in genes with diverse functions, including protein-glycosylation (*ALG14*), polyunsaturated fatty acid metabolism (*FADS1* and *FADS2*), phospholipid modeling (*LPGAT1*), and glucose- and oxygen- sensing pathways (*GCKR* and *HIF1AN*). These results expand our knowledge of genetic factors relevant to DNL and fatty acid biology, and will stimulate future research to investigate their potential role in cardio-metabolic diseases.

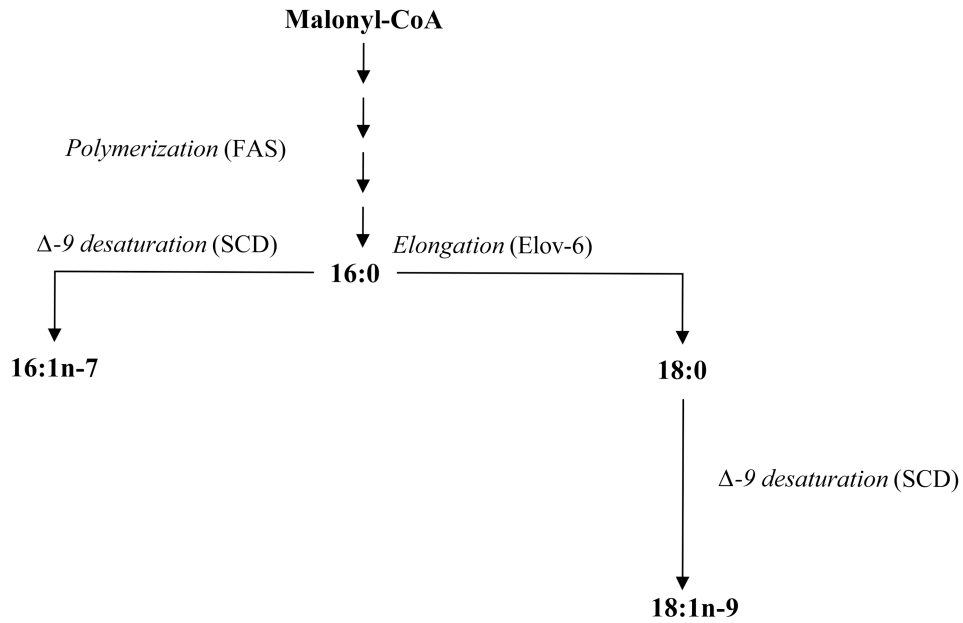


Figure 1.

Major fatty acids in the de novo lipogenesis pathway. Palmitic acid (16:0), stearic acid (18:0), palmitoleic acid (16:1n-7), and oleic acid (18:1n-9) are the major saturated and mono-unsaturated fatty acids that are obtained via de novo lipogenesis (DNL) or the diet, and investigated in this study (the metabolic steps which link the fatty acids are in italics, and the key enzyme involved in these processes labeled in brackets). In the initial steps of DNL, fatty acid synthase (FAS) catalyze the polymerization of malonyl-CoA to form 16:0 as the major initial product. 16:0 can be elongated to 18:0, and cellular studies suggested elongase-6 (Elov6) is the enzyme primarily responsible for this conversion²⁹. Both 16:0 and 18:0 are substrates for stearoyl-CoA desaturase (SCD) to give rise to 16:1n-7 and 18:1n-9, respectively².

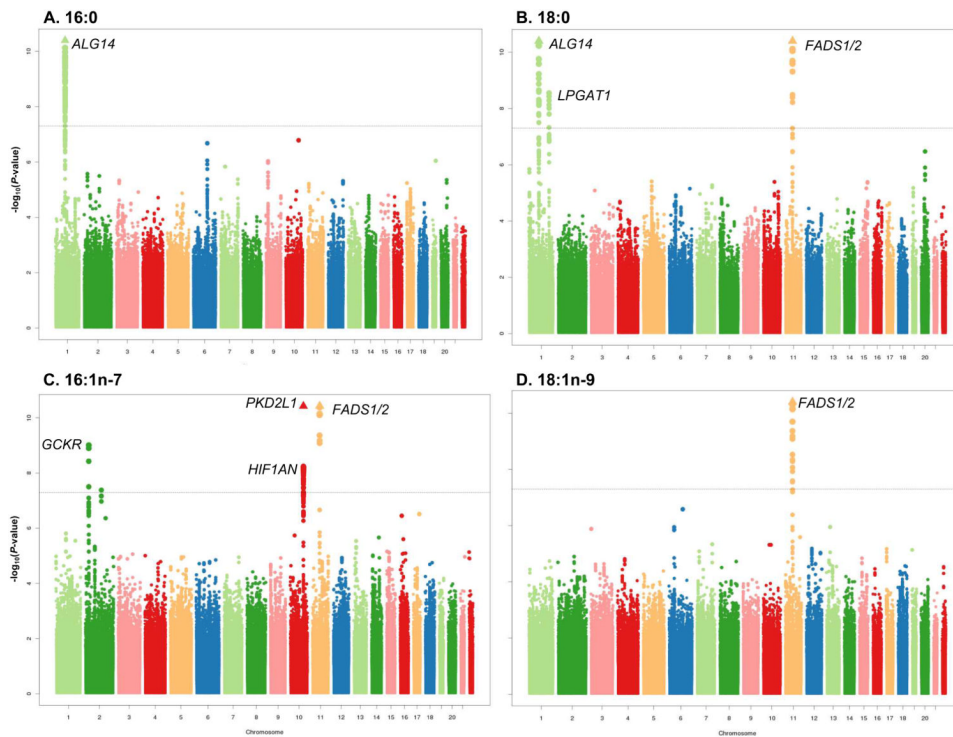


Figure 2. Meta-analysis of genome-wide associations with fatty acids in the de novo lipogenesis pathway. A: palmitic acid (16:0); B: stearic acid (18:0); C: palmitoleic acid (16:1n-7), and D: oleic acid (18:1n-9). Associations are demonstrated by chromosome location and $-\log_{10}$ (P -value), up to P -values of 10^{-10} . Triangles indicate additional SNPs with P -value $< 10^{-10}$. Genes of interest in each locus with SNPs variants which reached genome-wide significance are shown.

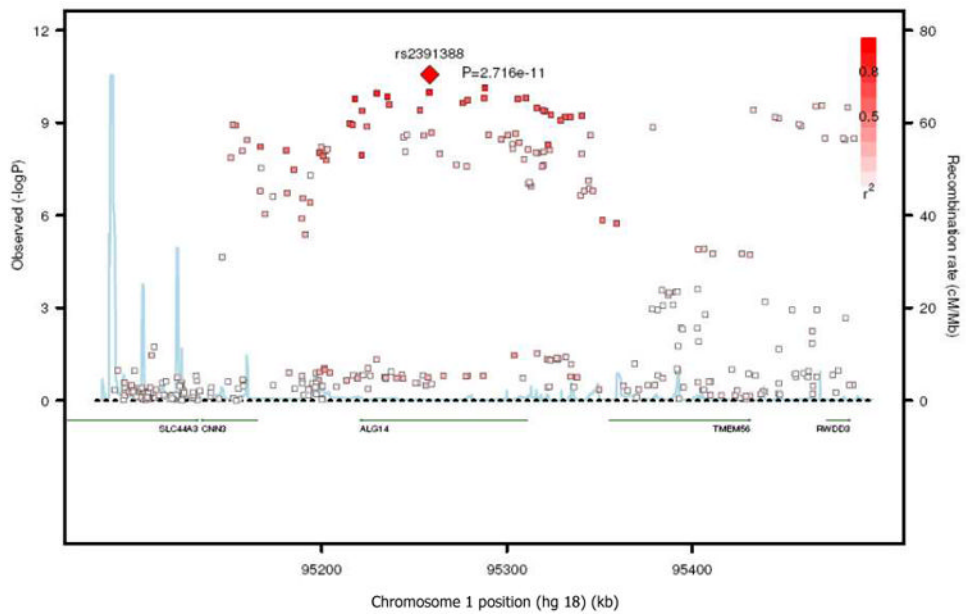


Figure 3. SNP association plots for palmitic acid-associated region. Genetic coordinates are along the x-axis (as per NCI build 36) and genome-wide association significance level is plotted against the y-axis as $-\log_{10}(P\text{-value})$. LD is indicated by color scale in relationship to marker rs2391388, with red for strong linkage disequilibrium (LD; $r^2 \geq 0.8$) and fading color for lower LD.

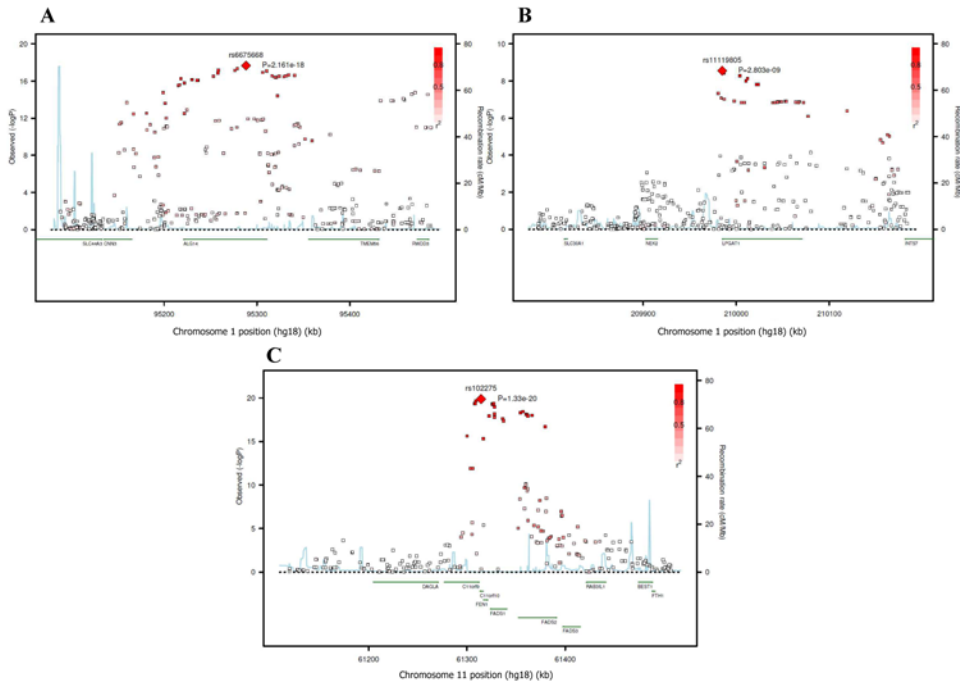


Figure 4. SNP association plots for stearic acid-associated region. Genetic coordinates are along the x-axis (as per NCI build 36) and genome-wide association significance level is plotted against the y-axis as $-\log_{10}(P\text{-value})$. (A) *ALG14* cluster region. LD is indicated by color scale in relationship to marker rs6675668. (B) *LPGAT1* cluster region. LD is indicated by color scale in relationship to marker rs11119805. (C) *FADS* cluster region. LD is indicated by color scale in relationship to marker rs102275. The color scheme is red for strong linkage disequilibrium (LD; $r^2 > 0.8$) and fading color for lower LD.

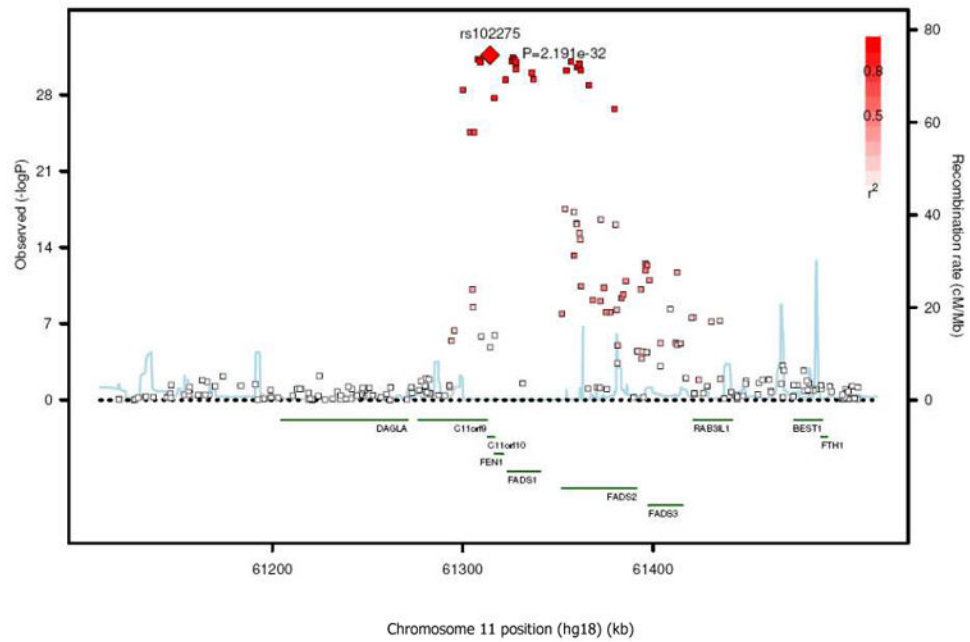


Figure 5. SNP association plots for oleic acid-associated region. Genetic coordinates are along the x-axis (as per NCI build 36) and genome-wide association significance level is plotted against the y-axis as $-\log_{10}(P\text{-value})$. LD is indicated by color scale in relationship to marker rs102275, with red for strong linkage disequilibrium (LD; $r^2 \geq 0.8$) and fading color for lower LD.

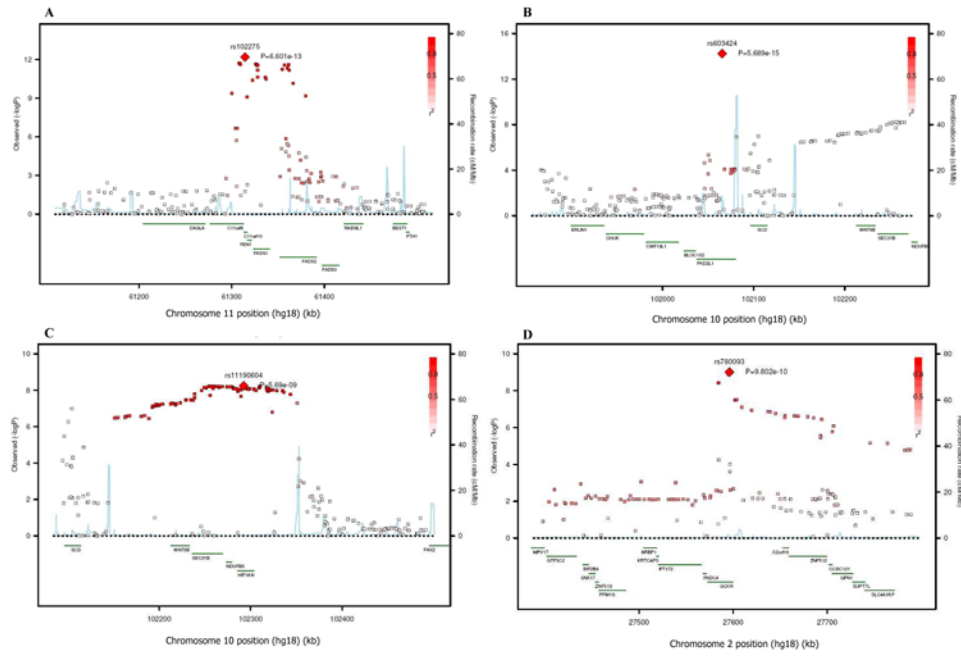
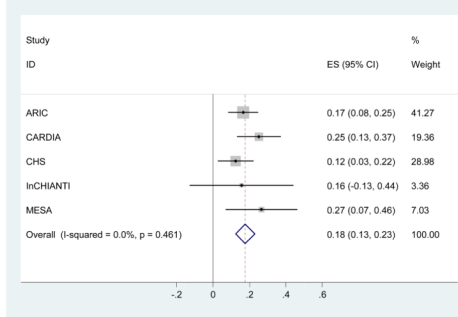
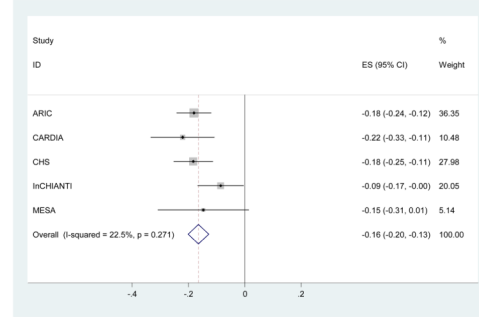


Figure 6. SNP association plots for palmitoleic acid-associated region. Genetic coordinates are along the x-axis (as per NCI build 36) and genome-wide association significance level is plotted against the y-axis as $-\log_{10}(P\text{-value})$. (A) *FADS* cluster region. LD is indicated by color scale in relationship to marker rs102275. (B) *PKD2L1* cluster region. LD is indicated by color scale in relationship to marker rs603424. (C) *HIF1AN* cluster region. LD is indicated by color scale in relationship to marker rs11190604. (D) *GCKR* cluster region. LD is indicated by color scale in relationship to marker rs780093. The color scheme is red for strong linkage disequilibrium (LD; $r^2 > 0.8$) and fading color for lower LD.

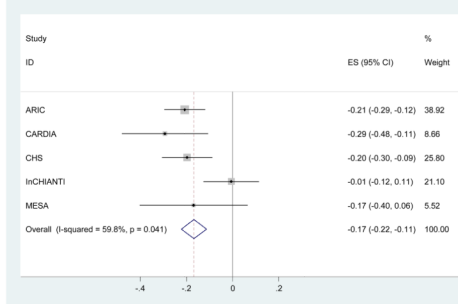
(1) 16:0 – rs2391388



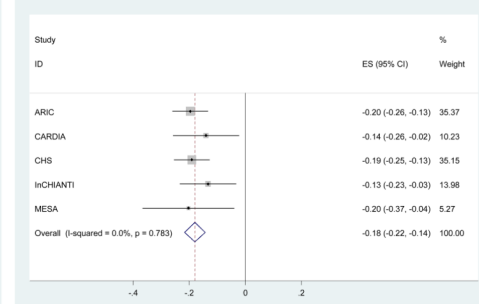
(2) 18:0 – rs6675668



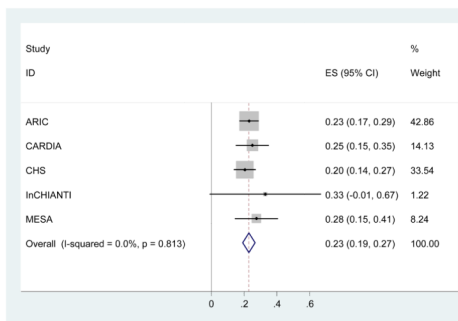
(3) 18:0 – rs11119805



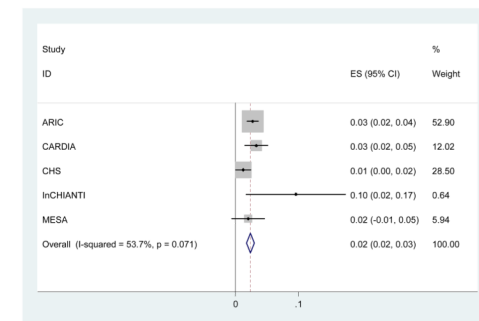
(4) 18:0 – rs102275



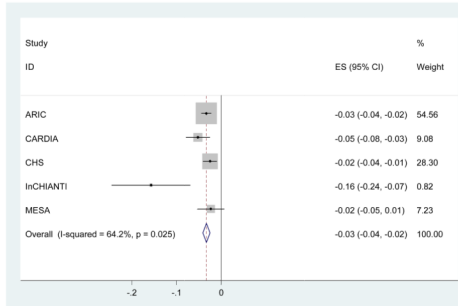
(5) 18:1n9 – rs102275



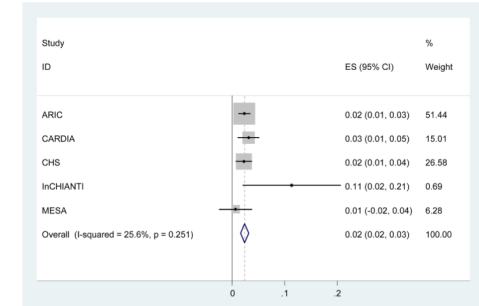
(6) 16:1n7 – rs102275

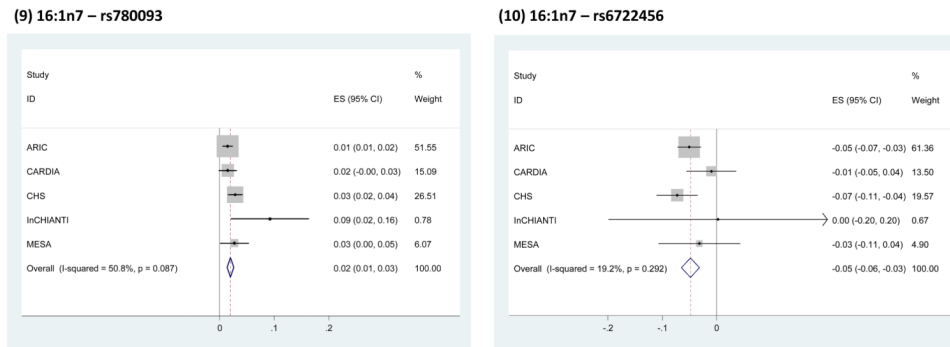


(7) 16:1n7 – rs603424



(8) 16:1n7 – rs11190604



**Figure 7.**

Forest plots for each of the top SNP-fatty acid associations: Within cohort effect size and 95% CI's were obtained from linear regression analysis using robust standard errors, and results were pooled using inverse-variance weighted meta-analysis. The size of the grey box around the central effect size estimate of each study is proportional to its inverse-variance weight in the meta-analysis. The vertical dashed line indicates the pooled meta-analysis effect size estimate. Chi-square test for heterogeneity *P*-values, and the *I*² statistic are also shown for each meta-analysis. The magnitude and direction of associations were generally consistent across all 5 cohorts. In cases where moderate heterogeneity were present, this was largely due to different findings in the InCHIANTI cohort. We note that all other cohorts assessed fatty acids in plasma phospholipids, whereas fatty acids were measured in total plasma in InCHIANTI; this difference could, at least in part, account for some of this heterogeneity. It is also important to note that this heterogeneity was due to differences in the magnitudes of the gene-fatty acid associations in InCHIANTI versus the other cohorts, rather than differing directions of associations; and that exclusion of InCHIANTI from each meta-analysis did not materially alter the top SNP-fatty acid associations (results not shown)

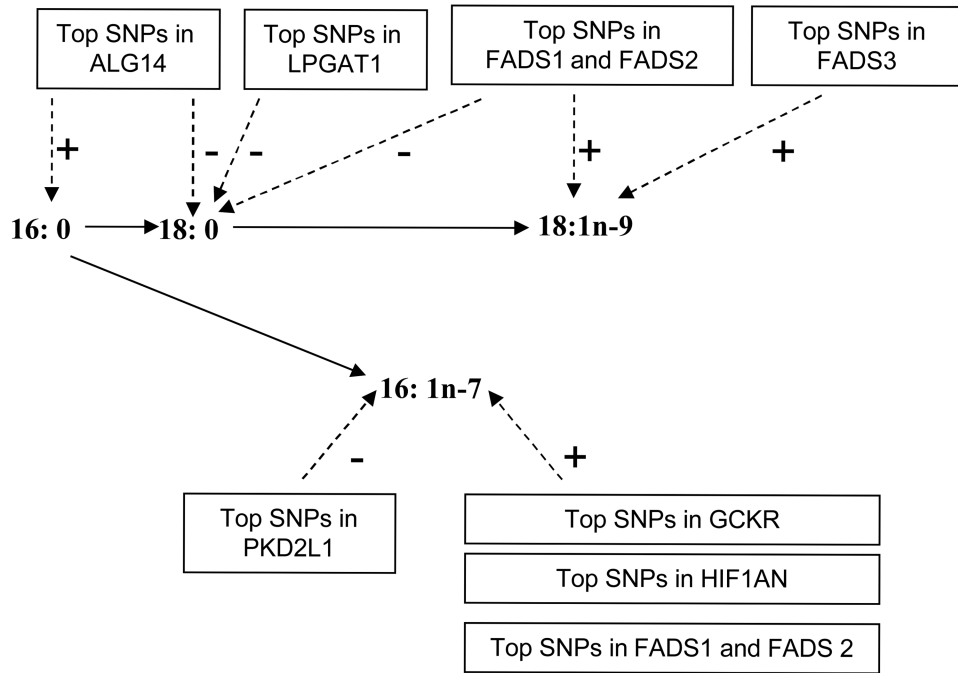


Figure 8. Summary of genome-wide association results. The genome-wide significant associations of identified loci (and genes of potential interest) with each fatty acid are shown with dashed arrows, and the +/- signs indicate the direction of the associations for the minor alleles at each loci.

Table 1

CHARGE cohort characteristics*

Cohort	n	Age, years	Men, %	BMI, kg/m ²	Fat intake, % of total energy [‡]	Carbohydrate intake, % of total energy [‡]	Alcohol intake, g/day [‡]	Plasma concentration, % of total fatty acids [‡]			
								16:0	16:1n-7	18:0	18:1n-9
ARIC	3269	53.8 (5.6)	48.7	27.0 (4.6)	34.0 (6.4)	46.5 (8.6)	8.2 (13.8)	25.4 (1.6)	0.64 (0.18)	13.3 (1.2)	8.6 (1.1)
CARDIA	1507	45.8 (3.4)	46.7	27.8 (5.9)	35.9 (8.7)	46.6 (9.1)	10.5 (19.8)	25.1 (1.8)	0.56 (0.22)	12.9 (1.6)	8.2 (1.2)
CHS	2404	72.0 (5.1)	38.4	26.4 (4.4)	32.2 (6.0)	53.8 (6.9)	5.2 (12.3)	25.5 (1.6)	0.51 (0.21)	13.4 (1.1)	7.6 (1.1)
InCHIANTI	1075	68.4 (15.5)	45.1	27.1 (4.2)	30.9 (5.1)	51.6 (6.8)	15.3 (20.9)	22.5 (2.4)	2.4 (0.9)	6.5 (1.0)	25.9 (3.7)
MESA	706	61.6 (10.4)	47	27.7 (5.2)	33.8 (7.1)	49.1 (8.5)	8.8 (14.6)	25.8 (1.9)	0.61 (0.25)	12.7 (1.6)	8.3 (1.2)

* Values are mean (SD) except where indicated otherwise.

[‡] Number of subjects in each cohort with dietary data were: ARIC, n=3192; CARDIA, n=1173; CHS, n=2404; InCHIANTI, n=1059 and MESA, n=658.

[‡] Fatty acids were measured in total plasma (InCHIANTI), and plasma phospholipid (all other cohorts).

Table 2
SNPs associated with plasma fatty acids in the DNL pathway at P values < 5.0 × 10⁻⁸

Most Significant SNP										
Fatty acid	Chr	Gene(s) of interest	Number of significant SNPs	rs number*	Position [†]	Coded allele	MAF	P	Beta coefficient (CI)	% of fatty acid variance explained [§]
16:0	1	<i>ALG14</i>	83	rs2391388	95258413	C	0.451	2.72 × 10 ⁻¹¹	0.178 (0.125, 0.23)	0.21-0.98
18:0	1	<i>ALG14</i>	98	rs6675668	95288225	T	0.49	2.16 × 10 ⁻¹⁸	-0.165 (-0.202, -0.128)	0.37-1.39
	1	<i>LPGATI</i>	8	rs11119805	209984867	A	0.123	2.8 × 10 ⁻⁹	-0.168 (-0.223, -0.113)	<0.01-0.72
	11	<i>FADS1/2</i>	33	rs102275	61314379	C	0.322	1.33 × 10 ⁻²⁰	-0.18 (-0.218, -0.142)	0.33-1.34
18:1n-9	11	<i>FADS1/2/3</i>	56	rs102275	61314379	C	0.328	2.19 × 10 ⁻³²	0.23 (0.192, 0.268)	0.32-2.14
16:1n-7	11	<i>FADS1/2</i>	23	rs102275	61314379	C	0.329	6.6 × 10 ⁻¹³	0.024 (0.017, 0.03)	0.15-1.03
	10	<i>PKD2LI</i>	1	rs603424	102065469	A	0.193	5.69 × 10 ⁻¹⁵	-0.033 (-0.041, -0.024)	0.28-1.57
	10	<i>HIF1AN</i>	70	rs11190604	102292447	G	0.218	5.69 × 10 ⁻⁹	0.024 (0.016, 0.032)	0.02-0.71
	2	<i>GCKR</i>	5	rs780093	27596107	T	0.41	9.8 × 10 ⁻¹⁰	0.02 (0.014, 0.027)	0.23-0.93
	2	-	2	rs6722456	134245561	A	0.023	4.12 × 10 ⁻⁸	-0.048 (-0.065, -0.031)	<0.01-0.57

* Loci are listed by fatty acids and in the order in which they are mentioned in the results section.

[†]The SNPs rs2391388 and rs6675668 are in LD ($r^2 = 0.875$).

[‡]Position for genome build 36.3.

[§] Calculated from the formula $(\beta^2 * 2 * \text{MAF} * (1 - \text{MAF})) / \text{Var}(Y)$, where β is the regression coefficient for one copy of the allele, MAF is the minor allele frequency and Var(Y) is the variance of the fatty acid. The values are the lowest and highest % variance explained across the 5 cohorts.