Prostaglandins, Leukotrienes and Essential Fatty Acids 94 (2015) 65–72

Contents lists available at ScienceDirect

Prostaglandins, Leukotrienes and Essential Fatty Acids

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

A genome-wide association study of saturated, mono- and polyunsaturated red blood cell fatty acids in the Framingham Heart Offspring Study ☆

N.L. Tintle [a,*], J.V. Pottala [b,c], S. Lacey [d], V. Ramachandran [e,f], J. Westra [a], A. Rogers [a], J. Clark [a], B. Olthoff [a], M. Larson [d,f,g], W. Harris [b,c,h], G.C. Shearer [i]

[a] Department of Mathematics, Statistics and Computer Science, Dordt College, Sioux Center, IA 51250, USA
[b] Health Diagnostic Laboratory, Richmond, VA, USA
[c] Department of Internal Medicine, Sanford School of Medicine, University of South Dakota, Sioux Falls, SD, USA
[d] Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Ave., Boston, MA, USA
[e] Framingham Heart Study, 73 Mt. Wayte Ave., Framingham, MA 01702, USA
[f] Boston University School of Medicine, 72 E. Concord St., Boston, MA 02118, USA
[g] Department of Mathematics and Statistics, Boston University, 111 Cummington St., Boston, MA, USA
[h] OmegaQuant, Sioux Falls, SD, USA
[i] Department of Nutritional Sciences, Pennsylvania State University, University Park, PA, USA

ARTICLE INFO

Article history:
Received 12 July 2014
Received in revised form
14 November 2014
Accepted 17 November 2014

Keywords:
PUFA
apoA-I maturation
Oleic acid
Linoleic acid
Arachidonic acid
Glycerol-lipids

ABSTRACT

Most genome-wide association studies have explored relationships between genetic variants and plasma phospholipid fatty acid proportions, but few have examined apparent genetic influences on the membrane fatty acid profile of red blood cells (RBC). Using RBC fatty acid data from the Framingham Offspring Study, we analyzed over 2.5 million single nucleotide polymorphisms (SNPs) for association with 14 RBC fatty acids identifying 191 different SNPs associated with at least 1 fatty acid. Significant associations ($p < 1 \times 10^{-8}$) were located within five distinct 1 MB regions. Of particular interest were novel associations between (1) arachidonic acid and *PCOLCE2* (regulates apoA-I maturation and modulates apoA-I levels), and (2) oleic and linoleic acid and *LPCAT3* (mediates the transfer of fatty acids between glycerolipids). We also replicated previously identified strong associations between SNPs in the *FADS* (chromosome 11) and *ELOVL* (chromosome 6) regions. Multiple SNPs explained 8–14% of the variation in 3 high abundance (> 11%) fatty acids, but only 1–3% in 4 low abundance (< 3%) fatty acids, with the notable exception of dihomo-gamma linolenic acid with 53% of variance explained by SNPs. Further studies are needed to determine the extent to which variations in these genes influence tissue fatty acid content and pathways modulated by fatty acids.

1. Introduction

Red blood cell (RBC) proportions of omega-3 and omega-6 fatty acids have well established relationships with a variety of disease phenotypes and risk factors, including total mortality [1], acute coronary syndrome [2,3], serum lipid levels [4], inflammatory markers [5], cognitive function [6] and brain size [7,8] among others. Variation in RBC omega-3 fatty acid levels (i.e., proportions, expressed as a percent of total fatty acids) have been shown to possess a strong heritable component [4,9], suggesting that not only dietary but also genetic factors likely play an important role in explaining differences between individuals [10,11].

Recently genome-wide association studies (GWAS) have sought to identify common single nucleotide polymorphisms (SNPs) with fatty acid levels. Initial investigations have focused on establishing potential relationships between plasma phospholipid fatty acid proportions and common SNP variations [12–15]. However, mounting evidence suggests that this fatty acid pool may be more affected by recent fat consumption [16], potentially obscuring the role of genetic variation in determining fatty acid composition [4,17].

Here we report a GWAS exploring relationships between the relative proportions of fourteen saturated, mono- and polyunsaturated RBC fatty acids with over 2.5 million common (minor allele frequency > 1%) SNPs in the Framingham Heart Offspring Study.

☆*Sources of support*: National Human Genome Research Institute (R15HG006915; NT), National Heart Lung and Blood Institute (NHLBI; R01 HL089590), NIH Contract N01-HC-25195, the Framingham Heart Study (NHLBI) and Boston University School of Medicine.

* Corresponding author. Tel.: +1 712 722 6264.
E-mail address: nathan.tintle@dordt.edu (N.L. Tintle).

http://dx.doi.org/10.1016/j.plefa.2014.11.007
0952-3278/

2. Materials and methods

2.1. Sample

Our analysis focused on the Framingham Heart Study (FHS) Offspring sample, a population based longitudinal study of families living in Framingham, Massachusetts. Detailed descriptions of the sample are available [4,18–20]. The final sample for this study consisted of 2633 individuals for whom both fatty acid and genotype data were available. The 2633 is a subset of 2899 Offspring subjects attending Examination 8 (2005–2008); those excluded due to missing genotype data had similar demographic and fatty acid profiles as those included in the study (data not shown). Written informed consent was provided by all participants, and the Institutional Review Board at the Boston University Medical Center approved the study protocol.

2.2. Fatty acid measurements

The fatty acid composition of RBC samples were analyzed by gas chromatography equipped with a SP 2560 capillary column after direct transesterification for 10 min in boron trifluoride/methanol and hexane at 100C as previously described [4]. This technique generates fatty acids primarily from RBC glycerophospholipids. All fatty acids with at least 0.5% abundance were included for analysis (except trans oleic acid): arachidonic acid (AA), dihomo-gamma-linoleneic acid (DGLA), docosahexaenoic acid (DHA), docosapentaenoic acid-n3 (DPA-n3), docosapentaenoic acid-n6 (DPA-n6), docosatetraenoic acid (DTA), eicosapentaenoic acid (EPA), linoleic acid (LA), oleic acid (OA), palmitic acid (PA), and stearic acid (SA). Three fatty acids below the 0.5% abundance level were also included: Palmitoleic acid (POA), because it is a marker of de novo lipogenesis [21]; gamma-linolenic acid (GLA), because it is the initial product of LA metabolism via delta-6 desaturase [22]; and alpha-linolenic acid (ALA), because as a dietary essential fatty acid like LA, its levels should be under less metabolic control than fatty acids that are the products of metabolism. Means and SDs of all 14 FAs are provided in Supplementary Table 1.

2.3. Genotype data

This analysis is based on approximately 2.5 million autosomal CEU (Centre d'Etude du Polymorphisme Humain collection from Utah; Northern and Western European ancestry) HapMap SNPs which were measured directly (approximately 550,000) or imputed (approximately 2 million) as previously reported [18]. Briefly, direct genotypes were obtained using the Affymetrix 500K and MIPS 50K chips, and were analyzed at the Affymetrix Core Laboratory. SNPs with missing data rates more than 3%, Hardy Weinberg Equilibrium p-values less than 1×10^{-6}, more than 100 Mendelian errors or low minor allele frequency (less than 1%) were eliminated from consideration. Measured SNPs, along with HapMap (release 22, build 36, CEU reference panel) [23] was used to impute the remaining 2 million SNPs using MaCH [24]. Quality control procedures were similar for both imputed and directly measured SNPs, with the addition of standard imputation quality metrics for imputed SNPs (see [18] for details).

2.4. Statistical analysis

For each of the 14 fatty acids considered here, a linear mixed model was fit for each of the 2.5 million SNPs which passed the initial quality control criteria. Each regression model predicted $\log_{10}$-transformed fatty acid level by SNP genotype (number of minor alleles), adjusting for age, sex and a random covariance component summarizing the family-structure present in the Framingham data set (i.e. matrix of kinship coefficients). We reported partial r^2 values in the text, which were percent change in the log of the relative fatty acid proportion explained by one additional minor allele (i.e. additive genetic model). Linear mixed effects models using the kinship matrix were run using *R* (lmekin function) for all analyses [25]. SNPs were considered genome-wide significant if their p-value was less than 1×10^{-8}. For each fatty acid, the distribution of p-values was evaluated using a Q–Q plot, and the genomic control lambda value (λ_{GC}) was estimated. λ_{GC} values ranged between 0.996 and 1.067 for the 8 fatty acids with at least one SNP reaching genome-wide significance ($p < 1 \times 10^{-8}$), showing little evidence for over-inflation of test statistics [26]. A final model for each fatty acid was also computed which adjusted for age, sex and kinship, as well as all significant SNPs, after eliminating SNPs showing strong pairwise correlations (Pearson correlation above 0.7 between genotypes) with other SNPs in the model.

3. Results

The clinical characteristics for the FHS Offspring participants and their fatty acid levels have been previously reported [4]; briefly they had a mean (SD) age of 66 (9) years, 54% were female, 9% smoked, and nearly half were treated for hypertension (49%) and high cholesterol (43%). They also had the following comorbidities: diabetes (14%), coronary heart disease (11%), and congestive heart failure (3%).

In all, 470 linear mixed models of SNP–fatty acid combinations reached a genome-wide significant p-value of less than 1×10^{-8} (Supplementary Table 2). Since some SNPs were related to multiple fatty acids, there were a total of 191 different (i.e., uniquely associated) SNPs reaching genome-wide significance for at least one fatty acid (see Supplementary Table 3). In general, when multiple fatty acids were associated with the same SNP, the fatty acids were correlated (details not shown), in part because they are often in the same metabolic pathway. Six of the fatty acids in our analysis had no SNPs reach the genome-wide significance level (DHA, DPA-n6, EPA, PA, POA and SA all$1 \times 10^{-8} < p\text{-value} < 1 \times 10^{-6}$), but eight of the fatty acids contained at least one genome-wide significant SNP (see Supplementary Fig. 1a–h for Manhattan plots of these eight fatty acids).

As expected, many of the 191 uniquely associated SNPs occur in close proximity to each other [i.e., within a 1 mega-base (MB) region], and so we summarized the 191 unique associations by focusing on five distinct 1 MB or smaller regions which contained all 191 uniquely associated SNPs. Table 1 provides an overview of

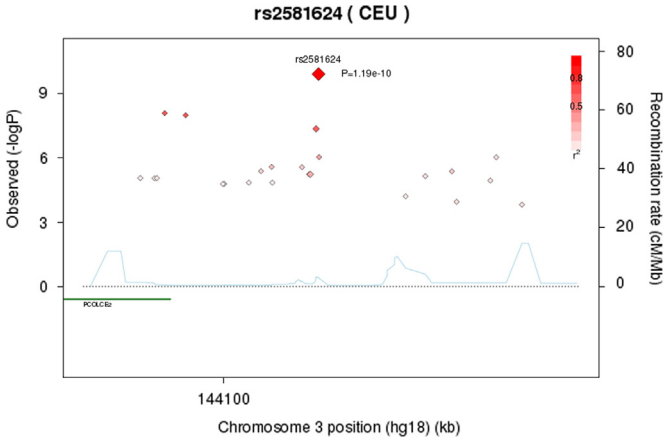

Fig. 1. Regional association plot of *PCOLCE2* locus with Arachidonic Acid (AA) levels. Correlations between the target SNP (rs2581624; the SNP with the lowest p-value (1.19×10^{-10})) and nearby SNPs (within a 500 kb) region are highlighted. Genotypes at rs2581624 are highly correlated with both rs2248811 (intergenic) and rs6778966 (intra-genic), both of which also reach genome-wide significance ($-\log P > 8$). r^2 values were based on the HapMap CEU population.