Atherosclerosis 235 (2014) 84–93

Contents lists available at ScienceDirect

Atherosclerosis

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

# Linkage analysis incorporating gene–age interactions identifies seven novel lipid loci: The Family Blood Pressure Program

Jeannette Simino [a,*,1], Rezart Kume [a,1,2], Aldi T. Kraja [b], Stephen T. Turner [c], Craig L. Hanis [d], Wayne H.-H. Sheu [e], Yii-Der Ida Chen [f], Cashell E. Jaquish [g], Richard S. Cooper [h], Aravinda Chakravarti [i], Thomas Quertermous [j], Eric Boerwinkle [d], Steven C. Hunt [k], D.C. Rao [a,l,m,n]

[a] Division of Biostatistics, Washington University in St. Louis, School of Medicine, 660 South Euclid Avenue, Campus Box 8067, Saint Louis, MO 63110-1093, USA
[b] Division of Statistical Genomics, Washington University in St. Louis, School of Medicine, Saint Louis, MO, USA
[c] Division of Nephrology and Hypertension, Mayo Clinic, Rochester, MN, USA
[d] Human Genetics Center, University of Texas Health Science Center, Houston, TX, USA
[e] Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan
[f] Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, Harbor-UCLA Medical Center, Torrance, CA 90502, USA
[g] Division of Cardiovascular Sciences, National Heart, Lung, Blood Institute, Bethesda, MD, USA
[h] Department of Preventive Medicine and Epidemiology, Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA
[i] McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD, USA
[j] Department of Medicine, Division of Cardiovascular Medicine, Cardiovascular Institute, Stanford University, Stanford, CA, USA
[k] Cardiovascular Genetics, Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA
[l] Department of Genetics, Washington University in St. Louis, School of Medicine, MO, USA
[m] Department of Psychiatry, Washington University in St. Louis, School of Medicine, MO, USA
[n] Department of Mathematics, Washington University in St. Louis, MO, USA

ARTICLE INFO

*Article history:*
Received 6 January 2014
Received in revised form
7 April 2014
Accepted 9 April 2014
Available online 26 April 2014

*Keywords:*
Lipids
Aging
Genetics
Meta-analysis
Gene–age interactions
Linkage
Triglyceride
Cholesterol

ABSTRACT

*Objective:* To detect novel loci with age-dependent effects on fasting (≥8 h) levels of total cholesterol, high-density lipoprotein, low-density lipoprotein, and triglycerides using 3600 African Americans, 1283 Asians, 3218 European Americans, and 2026 Mexican Americans from the Family Blood Pressure Program (FBPP).
*Methods:* Within each subgroup (defined by network, race, and sex), we employed stepwise linear regression (retention $p \leq 0.05$) to adjust lipid levels for age, age-squared, age-cubed, body-mass-index, current smoking status, current drinking status, field center, estrogen therapy (females only), as well as antidiabetic, antihypertensive, and antilipidemic medication use. For each trait, we pooled the standardized male and female residuals within each network and race and fit a generalized variance components model that incorporated gene–age interactions. We conducted FBPP-wide and race-specific meta-analyses by combining the *p*-values of each linkage marker across subgroups using a modified Fisher's method.
*Results:* We identified seven novel loci with age-dependent effects; four total cholesterol loci from the meta-analysis of Mexican Americans (on chromosomes 2q24.1, 4q21.21, 8q22.2, and 12p11.23) and three high-density lipoprotein loci from the meta-analysis of all FBPP subgroups (on chromosomes 1p12, 14q11.2, and 21q21.1). These loci lacked significant genome-wide linkage or association evidence in the literature and had logarithm of odds (LOD) score ≥ 3 in the meta-analysis with LOD ≥ 1 in at least two network and race subgroups (exclusively of non-European descent).
*Conclusion:* Incorporating gene–age interactions into the analysis of lipids using multi-ethnic cohorts can enhance gene discovery. These interaction loci can guide the selection of families for sequencing studies of lipid-associated variants.



* Corresponding author. Tel.: +1 314 362 8889; fax: +1 314 362 2693.
*E-mail address:* jeannette@wubios.wustl.edu (J. Simino).
[1] Authors contributed equally.
[2] The work was done while Mr. Kume was at the Division of Biostatistics, Washington University in St. Louis, School of Medicine, Saint Louis, Missouri, USA.

http://dx.doi.org/10.1016/j.atherosclerosis.2014.04.008
0021-9150/

## 1. Introduction

Lipid levels are important determinants of cardiovascular disease (CVD), an ailment that causes one of every three deaths in the United States and contributes to $312.6 billion in health care costs and lost productivity annually [1]. Although plasma lipid levels are highly heritable (40%–80%) [2,3], genetic association studies have identified (mostly) common variants in 157 loci that collectively explain less than 15% of the variance in lipid levels [4–6], leaving a substantial proportion of the lipid heritability missing [7]. The most recent (and largest) genetic association study of lipids to date included 188,577 participants of European descent and discovered 62 novel loci that collectively explained less than 2.6% of the lipid variance [5]. This insatiable need for increasingly bigger samples to detect common variants with diminishing effect sizes suggests that alternative approaches may facilitate the discovery of novel lipid loci with potentially larger effect sizes.

We leveraged the dynamic nature of genetic effects over a lifetime to enhance the discovery of lipid loci which may contain rare (minor allele frequency (MAF) $<$ 1%) and low-frequency (1% $\leq$ MAF $<$ 5%) variants. Plasma lipids vary by age [8] and exhibit greater correlations among first-degree relatives with similar ages [9], a promising but insufficient indicator of age-dependent genetic effects. Additional evidence has been rendered by candidate gene studies and a small-scale genome-wide association study (GWAS) which identified a handful of lipid loci with age-dependent effects [6,10-20]. Some of these were qualitative ("crossover") gene–age interactions, with opposite directions of the genetic effect in the young versus the elderly; for example, a single nucleotide polymorphism (SNP) in *SGSM2* was associated with increased low-density lipoprotein levels at ages 12–20 year old but decreased levels in 72–80 year olds [6]. Failing to account for gene–age interactions may prevent lipid loci from being detected, particularly in samples with wide age distributions. We incorporated gene–age interactions into genome-wide linkage analyses which may have a greater ability to detect low-frequency and rare variants than genome-wide association studies. Rare and low-frequency polymorphisms compose $\approx$74.6% of all variants [21] and may have larger effects sizes than common variants and explain substantial proportions of lipid variance. One recent study suggested that rare variants may collectively explain up to 7.8% of the variance in high-density lipoprotein levels [21], while another found that rare variants in four genes (*LPL, GCKR, APOB*, and *APOA5*) explained 1.1% of total variation in the diagnosis of hypertriglyceridemia [22].

In this investigation, our primary aim was to detect novel quantitative trait loci (QTLs) with age-dependent effects on fasting ($\geq$8 h) serum levels of total cholesterol (TC), high-density lipoprotein cholesterol (HDLC), low-density lipoprotein cholesterol (LDLC), and triglycerides (TG). We fit a highly flexible generalized variance components model [23], which allowed both the heritability and the correlations among relatives to depend on age [23], in each of nine population groups from the Family Blood Pressure Program (FBPP) [24]. These nine groups represented 3600 African Americans (AAs), 1283 Asians, 3218 European Americans (EAs), and 2026 Mexican Americans (MAs) [24]. We subsequently performed overall (across all races) and race-specific meta-analyses of these genome-wide linkage analyses to capitalize on the sharing of lipid loci across populations [3]. The newly discovered lipid loci may guide the selection of families to sequence using next-generation methods in order to increase the effect size and power to detect the causal variants underlying the linkage peak. This investigation may also provide further insight into the pervasiveness of age-dependent effects across the genome.

## 2. Methods

### 2.1. Subjects

The FBPP is comprised of four separate multicenter networks: GenNet, GENOA, HyperGEN, and SAPPHIRe. Although the recruitment method varied by the network, each recruited families ascertained through probands with high or low blood pressure. GenNet recruited nuclear families of young or middle-aged AAs, EAs (no lipid data available), and MAs through probands with untreated high-normal blood pressure. GENOA recruited AA and EA sibships containing at least two hypertensives, MA sibships containing at least 2 hypertensives who are also adult onset diabetics, and all full siblings of each index pair. HyperGEN recruited AA and EA sibships each containing at least two hypertensives, their parents, and one or more of a sibling's untreated offspring. SAPPHIRe recruited Chinese (CHI) and Japanese (JAP) sibpairs highly concordant or discordant for hypertension. More details on the participants are published elsewhere [24]. All FBPP studies were approved by their institutional review committees and all included participants provided informed consent.

### 2.2. Genotyping

The Mammalian Genotyping Service in Marshfield, Wisconsin, performed all genome-wide linkage scans [25]. For all cohorts except GenNet MAs, we used 362 autosomal markers that passed quality control criteria (had missing rate $<$70%, mapped to one location in the genome, and had a location not duplicated by other markers). GenNet MAs were genotyped later using a different screening set that contained 372 quality controlled autosomal markers; of which, 71% matched (by name and/or genetic distance) the autosomal markers from the other cohorts. The average inter-marker distances for the two screening sets were 9.5 cM and 9.4 cM, respectively. MapMaker/SIBS [26] and Pedcheck [27] were utilized to remove Mendelian errors, while GRR (graphical representation of relationships) [28] and ASPEX (affected sibpair exclusion mapping) [29] were used to verify family relationships.

### 2.3. Phenotype adjustment

We analyzed four fasting ($\geq$8 h) plasma lipid concentrations: TC, HDLC, LDLC, and the natural logarithm of TG (transformed to reduce the skewness and excess kurtosis). GenNet AAs and MAs had only one available lipid level, TC, measured by a finger stick (ProAct cholesterol system) [30]; GenNet EAs did not have any lipid information available. All racial groups from GENOA, HyperGEN, and SAPPHIRe had TC, HDLC, and TG measured by standard laboratory enzymatic methods. In general, LDLC was derived using the Friedewald equation [31] if TG $<$ 400 mg/dl and directly measured if TG $\geq$ 400 mg/dl. When multiple visit data were available, we included only one visit per individual (that which minimized missing lipid and covariate data). We excluded the lipid phenotypes of 87 participants with body-mass-index (BMI) or any lipid value that was four or more standard deviations away from the means of their subgroup (defined by network, race, and sex). Stepwise linear regression within each subgroup adjusted lipid phenotypes for age, age-squared, age-cubed, BMI, current smoking status (yes/no), current drinking status (yes/no), estrogen therapy (yes/no for females only) and field center (if multi-center), as well as dichotomous (yes/no) indicators of antidiabetic, antihypertensive, and antilipidemic medication use. We retained terms significant at the 0.05 level and standardized the residual lipid values to a mean of zero and standard deviation of one. We excluded any covariate-