Annals of human genetics

doi: 10.1111/ahg.12027

Post-Genome-Wide Association Study Challenges for Lipid Traits: Describing Age as a Modifier of Gene-Lipid Associations in the Population Architecture Using Genomics and Epidemiology (PAGE) Study

Logan Dumitrescu[1], Cara L. Carty[2], Nora Franceschini[3], Lucia A. Hindorff[4], Shelley A. Cole[5], Petra Bůžková[6], Fredrick R. Schumacher[7], Charles B. Eaton[8], Robert J. Goodloe[1], David J. Duggan[9], Jeff Haessler[2], Barbara Cochran[10], Brian E. Henderson[7], Iona Cheng[11], Karen C. Johnson[12], Chris S. Carlson[2], Shelly-Ann Love[3], Kristin Brown-Gentry[1], Alejandro Q. Nato Jr.[13], Miguel Quibrera[14], Garnet Anderson[2], Ralph V. Shohet[15], José Luis Ambite[16], Lynne R. Wilkens[17], Loïc Le Marchand[17], Christopher A. Haiman[7], Steven Buyske[18], Charles Kooperberg[2], Kari E. North[3,19], Myriam Fornage[20,21] and Dana C. Crawford[1,22]*

[1] *Center for Human Genetics Research, Vanderbilt University, Nashville, TN*
[2] *Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA*
[3] *Department of Epidemiology, University of North Carolina, Chapel Hill, NC*
[4] *Office of Population Genomics, National Human Genome Research Institute, Bethesda, MD*
[5] *Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX*
[6] *Department of Biostatistics, University of Washington, Seattle, WA*
[7] *Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA*
[8] *Department of Family Medicine, Warren Alpert Medical School, Brown University, Providence, RI*
[9] *Translational Genomics Research Institute, Phoenix, AZ*
[10] *Baylor College of Medicine, Houston, TX*
[11] *Cancer Prevention Institute of California, Fremont, CA*
[12] *Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN*
[13] *Department of Genetics, Rutgers University, Piscataway, NJ*
[14] *Gillings School of Public Health, University of North Carolina, Chapel Hill, NC*
[15] *John A. Burns School of Medicine, University of Hawaii, Honolulu, HI*
[16] *Information Sciences Institute, University of Southern California, Marina del Rey, CA*
[17] *Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI*
[18] *Department of Genetics and Department of Statistics, Rutgers University, Piscataway, NJ*
[19] *Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC*
[20] *Division of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas Health Sciences Center at Houston, TX*
[21] *Institute of Molecular Medicine, University of Texas Health Sciences Center at Houston, TX*
[22] *Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN*

Summary

Numerous common genetic variants that influence plasma high-density lipoprotein cholesterol, low-density lipoprotein cholesterol (LDL-C), and triglyceride distributions have been identified via genome-wide association studies (GWAS). However, whether or not these associations are age-dependent has largely been overlooked. We conducted an association study and meta-analysis in more than 22,000 European Americans between 49 previously identified GWAS variants

*Corresponding author: Dana C. Crawford, PhD, Center for Human Genetics Research, Vanderbilt University, 2215 Garland Avenue, 519 Light Hall, Nashville, Tennessee 37232–0700. Tel: 615 343 7852; Fax: 615 343 8619; E-mail: crawford@chgr.mc.vanderbilt.edu

and the three lipid traits, stratified by age (males: <50 or ≥50 years of age; females: pre- or postmenopausal). For each variant, a test of heterogeneity was performed between the two age strata and significant P_{het} values were used as evidence of age-specific genetic effects. We identified seven associations in females and eight in males that displayed suggestive heterogeneity by age ($P_{het} < 0.05$). The association between rs174547 (*FADS1)* and LDL-C in males displayed the most evidence for heterogeneity between age groups ($P_{het} = 1.74E-03$, $I^2 = 89.8$), with a significant association in older males ($P = 1.39E-06$) but not younger males ($P = 0.99$). However, none of the suggestive modifying effects survived adjustment for multiple testing, highlighting the challenges of identifying modifiers of modest SNP-trait associations despite large sample sizes.

Keywords: PAGE, modifier, age, lipids, genetic association

Introduction

Since 2005, 1617 genome-wide association studies (GWAS) have been published for common human diseases and traits yielding 249 genetic variants associated at genome-wide significance (Hindorff et al., 2011). A maturing step in the GWAS era is the identification of all common genetic variation associated with the disease or trait of interest. Arguably, this step is progressing rapidly for the lipid traits high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TG), for European-descent populations. The most recent meta-analysis of nearly 100,000 European-descent populations identified 95 loci associated with HDL-C, LDL-C, and/or TG (Teslovich et al., 2010).

Beyond identification of all common genetic variants, there is much interest in the post-GWAS era in identifying modifiers of GWAS-identified variants (Ober & Vercelli, 2011). Sex, age, and environmental exposures such as diet and smoking status are known to affect lipid levels (Third Report of the National Cholesterol Education Program (NCEP) Expert Panel, 2002), but little is known about their interactions with genetic variants associated with lipid levels. The understanding of these potential modifiers of genetic effects may be important in both public health genomics and personalized medicine as, unlike one's genetic make-up, some environmental exposures have the potential to be attenuated or amplified. Furthermore, understanding of these modifiers may also lead to a better understanding of the biology of the disease process, perhaps leading to improved preventive and/or treatment approaches (Manolio et al., 2006; Ober & Vercelli, 2011).

As mentioned above, it is well known that lipid levels can vary considerably with age, with mean total cholesterol levels tending to increase with increasing age (Reilly et al., 1990; Ericsson et al., 1991; Boomsma et al., 1996; Snieder et al., 1997). But the change is not always gradual. Indeed, there are four specific times in the life cycle that result in significant changes in one's lipid profile: birth and the first few following years, adolescence, middle age, and advanced age (Snieder et al., 1997). Often, these age-related changes differ among the sexes. For example, total cholesterol levels tend to peak around age 50 in men (Kronmal et al., 1993; Jousilahti et al., 1996) and at menopause in women (Hjortland et al., 1976; Bonithon-Kopp et al., 1989; Matthews et al., 1989). Why lipid levels change over time is likely due to a number of factors, but whether or not the genetic effects of variants important in lipid metabolism vary with age has yet to be thoroughly explored.

While the interest in and importance of identifying modifiers of genetic effects, such as age, is recognized, so too are the challenges in designing a study to identify such modifiers. The Population Architecture using Genomics and Epidemiology (PAGE) study, a collaboration of investigators with access to large population-based cohorts and cross-sectional epidemiologic studies, began in 2008 with the aim of characterizing GWAS-identified variants, and their modifiers, for a number of different traits (Matise et al., 2011). In one report, we replicated and characterized 49 GWAS-identified lipid associations in the diverse population-based samples available within the PAGE study (Dumitrescu et al., 2011). In this report, we extend our initial post-GWAS work to investigate if age modifies the association of these GWAS-identified variants for HDL-C, LDL-C, and TG within ~22,000 European-descent participants.

Methods

Study Populations and Phenotypes

Briefly, PAGE study samples were drawn from four large population-based studies or consortia: Epidemiologic Architecture for Genes Linked to Environment (EAGLE), based on three National Health and Nutrition Examination Surveys (NHANES; Centers for Disease Control and Prevention [CDC], 2002; CDC, 2004; CDC, 2010), the Multiethnic Cohort (MEC; Kolonel et al., 2004), the Women's Health Initiative (WHI; The Women's Health Initiative Study Group,

1998; Anderson et al., 2003), and Causal Variants Across the Life Course (CALiCo), a consortium of several cohort studies: Atherosclerosis Risk in Communities Study (ARIC; The ARIC Investigators, 1989), Coronary Artery Risk Development in Young Adults (CARDIA; Friedman et al., 1988), and Cardiovascular Health Study (CHS; Fried et al., 1991). All selected individuals are of European-descent and all studies were approved by Institutional Review Boards at their respective sites.

While the overall PAGE study consists of individuals representing most decades of the human age spectrum, it is important to note that each individual PAGE study site collected samples at different stages in the life course. Several PAGE study sites represent older cohorts (CHS, MEC, and WHI) and younger- to middle-aged cohorts (ARIC and CARDIA). Only one study, EAGLE accessing NHANES, collected data for all age ranges (from 12 years of age to the elderly). For studies with repeated measures, baseline age and lipid levels were used in this analysis.

Serum HDL-C, TG, and total cholesterol were measured using standard enzymatic methods. LDL-C was calculated using the Friedewald equation (Arking et al., 2006), with missing values assigned for samples with TG levels greater than 400 mg/dl. For PAGE study sites with longitudinal data, the baseline measurement was used for analysis.

SNP Selection and Genotyping

SNP selection, genotyping, and quality control have been previously described (Dumitrescu et al., 2011). Briefly, each PAGE study site performed genotyping independently using commercially available genotyping arrays (Affymetrix 6.0, Illumina 370CNV BeadChip), custom mid- and low-throughput assays (TaqMan, Sequenom, Illumina GoldenGate or BeadXpress), or a combination thereof (Affymetrix, Santa Clara, CA, USA; Illumina, San Diego, CA, USA; TaqMan, Applied Biosystems, Foster City, CA, USA; Sequenom, San Diego, CA, USA). Quality control was implemented at each study site using common criteria. A total of 49 SNPs reported here were genotyped in two or more PAGE study sites and were previously associated with HDL-C, LDL-C, and/or TG in published (as of 2008) candidate gene and GWAS. Characteristics for all 49 SNPs are listed in Table S1.

Statistical Methods

All tests of association were performed by each PAGE study site using a prespecified analysis protocol prior to meta-analysis. The study protocol was limited to fasting participants of European-descent regardless of lipid lowering medication usage given that previous analyses demonstrated that lipid lowering medication usage did not appreciably alter single-SNP association results, as treated individuals represented only a small subset of the overall study (Dumitrescu et al., 2011). Participants with >1000 mg/dl were excluded from the TG analyses, as very high TG levels may be due to rare mutations, disease or illness, or simple human error when recording the data. TG levels were natural-log transformed.

Linear regression was performed with HDL-C, LDL-C, or ln(TG) as the dependent variable and an SNP as the independent variable, assuming an additive genetic model. Sex effects on lipid trait distributions are well known (Reilly et al., 1990; Third Report of the National Cholesterol Education Program (NCEP) Expert Panel, 2002), and one PAGE study site (WHI) is composed of only women. Therefore, analyses were stratified by sex unless otherwise indicated. Selected PAGE study sites also included study site or site of ascertainment as a covariate in all models. Age distributions varied across the different PAGE studies (Tables S1-S3), with little age overlap across some studies. Therefore, to gain the largest sample size possible for each analysis, individuals were stratified into two age groups. Males were divided into two categories based on age (<50 years or ≥50 years of age), while females were divided into two categories based on menopausal status (premenopausal and postmenopausal). Participants <18 years of age and >70 years of age were excluded.

Meta-analysis, using a fixed effects inverse-variance-weighted approach, was first performed across PAGE studies within each age group using Meta-Analysis Helper (METAL; Willer et al., 2010). Then, for each variant, a 1-degree of freedom χ^2 test for heterogeneity was performed between the two age groups. The quantity I^2, a measure of the degree of inconsistency between age groups, was calculated using the equation $I^2 = 100\% \times (Q \text{ degrees of freedom})/Q$ (Higgins et al., 2003). Genetic associations were considered modified by age if the test of heterogeneity between age strata was significant at $P_{het} < 0.05$, uncorrected for multiple testing. Aggregate data from the meta-analysis as well as individual tests of association from each PAGE study site will be made available via dbGaP (Mailman et al., 2007; Matise et al., 2011).

Results

Table 1 summarizes the characteristics of the PAGE samples studied here, including the total number of individuals, mean age, and mean lipid levels for each age group in both males and females. Out of a total of 8550 male participants and 13,772 female participants, the majority fell into the older age group (68% of males were 50+ years of age and 76% of females were postmenopausal), reflecting the study recruitment of PAGE studies. The older age group was nearly 30 years older on average than the younger age group for both sexes. For

Table 1 PAGE Study Characteristics.

	Males		Females	
Trait	Age <50 years	Age ≥50 years	Premenopausal	Postmenopausal
N	2727	5823	3335	10,437
Age (years)	35.7 (10.3)	63.2 (9.0)	38.3 (11.8)	66.0 (8.6)
HDL-C (mg/dl)	44.3 (11.5)	46.8 (14.3)	56.1 (14.8)	57.3 (16.4)
LDL-C (mg/dl)	122.4 (35.3)	135.4 (35.6)	114.1 (33.9)	134.3 (36.7)
TG (mg/dl)	117.9 (96.1)	144.0 (83.5)	91.5 (58.1)	149.6 (81.2)

All values reported as mean (SD) unless otherwise indicated. All individuals are of European-descent.

all three lipid traits and both sexes, lipid levels were slightly higher in the older age group than in the younger age group (Table 1). Mean lipid levels by each PAGE study site are shown in Tables S2-S4.

For each variant, a test of heterogeneity between age groups was used to identify associations where age may modify the effect of the SNP on lipid levels. In meta-analyses stratified by sex and menopausal status, seven tests of heterogeneity were significant in females at $P_{het} < 0.05$: two for HDL-C, two for LDL-C, and three for (ln)TG (Table 2). The level of heterogeneity for these tests was high, ranging from 74.3% to 87.9%. The most significant test of heterogeneity in females was rs9989419 at *CETP* for LDL-C ($P_{het} = 3.98E-03$; $I^2 = 87.9$). This association was significant in postmenopausal women but not in premenopausal women and the genetic effect size (β) was in opposite directions among the two age groups ($P = 0.32$, $\beta = -0.84$ and $P = 1.25E-03$, $\beta = 2.36$ in pre- and postmenopausal women, respectively). Indeed, a majority (4/7) of the significant tests of heterogeneity involved an SNP-lipid trait association that was significant in one age group (at $P < 0.05$) but not the other. However, not all the significant tests of heterogeneity were as easily interpretable: more than half (4/7) of the genetic effect sizes in pre- versus postmenopausal females trended in the same direction, two involved an SNP-lipid trait association that was significant in both age groups, and one SNP-lipid trait association was not significant in either strata. Interestingly, for SNPs with associations stronger among premenopausal women compared with postmenopausal women, such as rs3764261 and rs4803750 (Table 2), these data are consistent with the hypothesis that genetics plays a greater role early in the aging process compared with the environment.

Tests of heterogeneity suggested that eight genetic associations were modified by age (<50 vs. ≥50 years) in males: three for HDL-C, three for LDL-C, and two for (ln)TG ($P_{het} < 0.05$; Table 3). The level of heterogeneity for these tests was high, ranging from 74.9% to 89.8%. None of the significant heterogeneity tests identified in males overlapped with those identified for females (Table 3 vs. Table 2). One variant, rs6102059 at *MAFB*, is listed twice in Table 3 as its associations with both HDL-C and ln(TG) displayed evidence of heterogeneity. The most significant test of heterogeneity in males was for *LIPC* rs174547 and LDL-C ($P_{het} = 1.74E-03$; $I^2 = 89.8$). As in females, this significant test of heterogeneity appears to be driven by the fact that the SNP-trait association was highly significant in one age stratum but not the other ($P = 0.99$ and $P = 1.39E-06$ in males <50 and ≥50 years old, respectively). Unlike in females, however, this trend was observed for all eight significant tests of heterogeneity.

Given that these tests of association stratified by sex and age were greatly impacted by sample size and power, we also performed tests of association stratified by age with sexes combined (Table 4). Premenopausal women were analyzed together with men <50 years of age, while postmenopausal women were analyzed together with men >50 years of age. In this sex-combined analysis, eight tests of heterogeneity were significant at $P_{het} < 0.05$: two for HDL-C, four for LDL-C, and two for (ln)TG. A majority (5/8) of the significant heterogeneity tests identified in all participants overlapped with those identified in females or males only (Table S5). For these overlapping associations, the magnitude of heterogeneity (I^2) was, in general, slightly lower in the sex-combined analysis than the sex-stratified analysis. The most significant test of heterogeneity in all participants was rs4803750 at *BCL3* for LDL-C ($P_{het} = 5.88E-03$; $I^2 = 86.8$), which was also significant in the female-only analysis. Associations that showed evidence of modification by age in the sex-combined analysis only included rs1699614 (*CILP2/PBX4/NCAN*) for HDL-C, rs6586891 (*LPL*) for HDL-C, and rs11206510 *(PCSK9*) for LDL-C.

Discussion

Epidemiologic studies have long documented that lipid trait distributions are shaped by sex, age, and environmental exposures such as smoking or diet. As lipid levels tend to increase with age, we explored the potential modifying effect of age on previously identified lipid-SNP associations by testing for heterogeneity between single-SNP associations stratified by

Table 2 Lipid SNPs with evidence of menopausal status as a modifier in females of European-descent.

				Premenopausal			Postmenopausal			All females				
SNP	Gene	Lipid trait	CA (CAF)	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	P_{het}	I^2
rs2197089	*LPL*	HDL	T (0.54)	0.29 (0.36)	0.41	3291	1.20 (0.29)	3.17E-05	8123	0.84 (0.24)	1.73E-04	11,414	0.049	74.3
rs3764261	*CETP*	HDL	T (0.33)	4.48 (0.37)	1.97E-33	3201	3.41 (0.38)	1.09E-19	4677	3.95 (0.26)	1.58E-50	7878	0.042	75.7
rs4803750	*BCL3*	LDL	A (0.93)	13.08 (2.52)	2.04E-07	1473	5.28 (1.65)	1.36E-03	4536	7.62 (0.38)	3.32E-08	6009	9.42E-03	85.2
rs9989419	*CETP*	LDL	A (0.39)	−0.84 (0.83)	0.32	3041	2.36 (0.73)	1.25E-03	5239	0.97 (0.55)	0.08	8280	3.98E-03	87.9
rs1883025	*ABCA1*	TG	T (0.26)	−0.05 (0.02)	1.73E-03	1111	−0.01 (0.02)	0.69	1166	−0.03 (0.01)	0.02	2277	0.038	76.8
rs562338	*APOB*	TG	T (0.20)	0.02 (0.01)	0.27	3192	−0.03 (0.01)	0.02	7084	−0.01 (0.01)	0.24	10,276	0.025	80.0
rs754523	*APOB*	TG	T (0.69)	0.02 (0.01)	0.13	3176	−0.02 (0.01)	0.07	4777	−0.003 (0.01)	0.67	7953	0.021	81.3

Single-SNP tests of association were performed within each age group and in the combined dataset. Beta and standard error (SE), P-value (P), and sample size (N) are given for each association. The beta estimate is per additional copy of the coded allele (CA). Tests of heterogeneity were performed between the two age strata and the resulting P-value (P_{het}) and I^2 index are also reported. Only results where $P_{het} < 0.05$ are reported.
CAF, coded allele frequency.

Table 3 Lipid SNPs with evidence of age as a modifier in males of European-descent.

				Males <50			Males ≥50			All males				
SNP	Gene	Lipid trait	CA (CAF)	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	P_{het}	I^2
rs3135506	*APOA1/C3/A4/A5*	HDL	C (0.06)	−0.41 (0.80)	0.61	1677	−2.42 (0.54)	6.26E-06	5325	−1.80 (0.45)	5.38E-05	7002	0.037	77.1
rs4775041	*LIPC*	HDL	C (0.29)	0.61 (0.34)	0.07	2674	1.52 (0.28)	5.47E-08	5438	1.15 (0.22)	9.48E-08	8112	0.038	76.7
rs6102059	*MAFB*	HDL	T (0.30)	1.08 (0.44)	0.01	1695	−0.04 (0.32)	0.91	3614	0.35 (0.26)	0.17	5309	0.041	76.2
rs174547	*FADS1*	LDL	T (0.67)	0.01 (1.03)	0.99	2391	4.26 (0.88)	1.39E-06	3452	2.46 (0.67)	2.41E-04	5843	1.74E-03	89.8
rs2156552	*LIPG*	LDL	A (0.84)	3.11 (1.25)	0.01	2404	−0.26 (0.91)	0.77	5324	0.91 (0.73)	0.21	7728	0.029	79.0
rs1566439	*CETP*	LDL	T (0.60)	2.71 (1.37)	0.05	1424	−0.38 (0.72)	0.60	4832	0.30 (0.64)	0.64	6256	0.046	74.9
rs1748195	*ANGPTL3*	TG	C (0.67)	0.09 (0.02)	1.47E-05	1552	0.02 (0.01)	0.11	3460	0.04 (0.02)	2.48E-04	5012	4.89E-03	87.4
rs6102059	*MAFB*	TG	T (0.30)	−0.05 (0.02)	0.02	1537	0.003 (0.01)	0.82	3517	−0.01 (0.01)	0.31	5054	0.032	78.3

Single-SNP tests of association were performed within each age group and in the combined dataset. Beta and standard error (SE), P-value (P), and sample size (N) are given for each association. The beta estimate is per additional copy of the coded allele (CA). Tests of heterogeneity were performed between the two age strata and the resulting P-value (P_{het}) and I^2 index are also reported. Only results where $P_{het} < 0.05$ are reported.
CAF, coded allele frequency.

Table 4 Lipid SNPs with evidence of age as a modifier in all individuals of European-descent.

				Younger males & females			Older males & females			All			
SNP	Gene	Lipid trait	CA	Beta (SE)	P	N	Beta (SE)	P	N	Beta (SE)	P	N	P_{het}	I^2
rs1699614	*CILP2/PBX4/NCAN*	HDL	T (0.12)	0.69 (0.43)	0.11	5965	−0.49 (0.34)	0.16	13,509	−0.03 (0.27)	0.92	19,474	0.033	77.9
rs6586891	*LPL*	HDL	A (0.64)	0.57 (0.25)	0.02	5987	1.24 (0.20)	8.06E-10	13,579	0.97 (0.16)	5.31E-10	19,566	0.034	77.7
rs11206510	*PCSK9*	LDL	T (0.79)	3.27 (0.79)	3.44E-05	5487	0.92 (0.61)	0.13	12,349	1.81 (0.48)	1.86E-4	17,836	0.019	81.8
rs174547	*FADS1*	LDL	T (0.67)	1.15 (0.66)	0.08	5442	3.19 (0.69)	3.32E-06	6338	2.13 (0.48)	7.50E-06	11,780	0.032	78.3
rs4803750	*BCL3*	LDL	A (0.93)	11.10 (2.03)	4.38E-08	2393	4.74 (1.10)	1.66E-05	9023	6.19 (0.97)	1.58E-10	11,416	5.88E-03	86.8
rs9989419	*CETP*	LDL	A (0.40)	−0.29 (0.63)	0.64	5408	1.40 (0.51)	5.83E-03	10,433	0.74 (0.40)	0.06	15,841	0.036	77.1
rs1748195	*ANGPTL3*	TG	C (0.67)	0.06 (0.01)	3.85E-06	3597	0.02 (0.01)	0.06	7926	0.03 (0.01)	1.72E-05	11,523	0.011	84.5
rs1883025	*ABCA1*	TG	T (0.26)	−0.04 (0.01)	8.21E-03	2065	−0.003 (0.01)	0.81	2669	−0.01 (0.01)	0.07	4734	0.049	74.1

Single-SNP tests of association were performed within each age group and in the combined dataset. Beta and standard error (SE), P-value (P), and sample size (N) are given for each association. The beta estimate is per additional copy of the coded allele (CA). Tests of heterogeneity were performed between the two age strata and the resulting P-value (P_{het}) and I^2 index are also reported. Only results where $P_{het} < 0.05$ are reported.
CAF, coded allele frequency.

age or menopausal status. At a nominal $P_{het} < 0.05$, seven associations in females, eight associations in males, and eight associations in the sex-combined dataset displayed significant heterogeneity by age. In total, after accounting for overlap of results among the sex-stratified and sex-combined analyses, 18 distinct associations (comprising 17 SNPs and all three lipid traits) showed suggestive modifying effects of age.

It should be noted, however, that none of the P-values presented here were adjusted for multiple testing. Indeed, if we were to use a Bonferroni adjustment for 441 tests of heterogeneity (= 49 SNPs × 3 lipid traits × 3 sex-defined datasets), none of our tests would remain significant at P < 1.34E-04 (= 0.05/441). Of course, these associations are highly correlated and correction for 441 tests is overly conservative. To further complicate interpretation, tests of heterogeneity are generally known to be underpowered when only a small number of studies is included in the meta-analysis (Hardy & Thompson, 1998). Due to potentially low power, some investigators have recommended using a non-Bonferroni-adjusted significance threshold of $P_{het} < 0.10$, rather than the standard $P_{het} < 0.05$ (Higgins et al., 2003). Given this limitation, our results may err on being overly conservative. Furthermore, the high I^2 values observed here provide additional evidence of heterogeneity of genetic effects between age groups, with up to nearly 90% of the total variability in beta estimates attributed to heterogeneity. Therefore, while replication remains the gold standard, we present our most significant findings in the interest of hypothesis generation.

As potential environmental modifiers are often ignored in genetic association studies, it has been argued that accounting for these modifiers may help with replication across studies and with generalization across racial/ethnic groups. Using the association between rs1883025 at *ABCA1* and TG in the sex-combined analysis (Table 4) as an example, when all ages are included the resulting P-value is suggestive of association (P = 0.07) but would not meet most studies' statistical significance threshold. However, when this association is stratified into younger and older age groups, the SNP effect is clearly present in the younger cohort ($\beta = -0.04$ and P = 8.21E-03) but not the older cohort ($\beta = -0.003$ and P = 0.81), even though the sample sizes are similar.

To our knowledge, only a few studies have explored whether age modifies lipid-SNP associations. For example, Shirts et al. (2011) tested 17 previously identified lipid-SNP associations for interactions with age. No evidence for modification was found for the majority of associations; however, a single significant gene-age interaction (rs646776 at *SORT1* and LDL-C) was identified and replicated (Shirts et al., 2011). We found no evidence of age-related modification of this association in our data. Neither did our results support those in Webster et al. (2009), which showed some evidence of an age effect for rs328 and HDL-C and rs3135506 and TG.

Again, we found no evidence that these SNPs are involved in age-related changes in lipid levels. It is important to note that, in general, these studies found little evidence that the effects of previously established SNP-lipid associations vary significantly over time, a conclusion concordant with our results.

Our study faced many of the challenges inherent in studies of gene-environment or gene-gene interactions (Hunter, 2005; Ober & Vercelli, 2011). As mentioned earlier, correction for multiple testing can be a major burden that may only be overcome by large effect sizes or large population sizes. PAGE is a large and diverse collaboration. However, our study was limited to individuals of European-descent given that, following stratification by sex and age group, our power would be greatly limited in other racial/ethnic groups with smaller sample sizes. Therefore, the results presented here are not necessarily generalizable to populations of non-European-descent.

Another challenge commonly faced by genetic studies of environmental modifiers is interpretation of results. Even when a statistically significant modifier effect is identified, it is often difficult to dissect the biological etiology of the effect. This study was no exception. While age is an easy variable to collect and harmonize across studies, it is most likely a surrogate for numerous accumulating physiological changes. Indeed, there are many possibilities as to why we observe age-related heterogeneity, including survival bias, discrepant medication use between age groups (e.g., hormone replacement therapy in postmenopausal women), and genetic differences between younger and older populations not examined here. Also, it is difficult to distinguish between the effects of age and sex, given that both are surrogates for age- and sex-dependent exposures. In this analysis, we did not explicitly test for sex as a modifier of lipid trait profiles given its complex relationship with age. Despite our attempts to account for both variables in stratified analyses, the exact mechanism by which age modifies genetic effects on lipid levels is difficult to elucidate without further experimentation.

The lipid levels tested in this study were based on only one time point for each individual. Therefore, we were unable to compare interindividual change in lipid levels and their associations with genetic variants over time. While longitudinal studies are ideal for studying any environmental modifier, they are more expensive, tend to be smaller in sample size, and require more time to collect exposure data compared to cross-sectional studies. Therefore, cross-sectional and case-control studies may be used as a first pass for identifying potential modifiers. Additionally, we must note that our analysis was limited to 49 established loci drawn from GWAS studies published in or before 2008. Therefore, as additional GWAS have been performed, many additional lipid-associated loci have been identified, which we were not able to investigate here.

The results of this study have an important implication for genetic association studies of lipid traits. While it is common practice to simply adjust for age as a confounder, this assumes that the relationship between age and lipid levels is homogenous across the different genotypes of the variant under study. Our results suggest that this assumption does not hold true for some variants. Therefore, caution must be exercised when combining or comparing genetic associations across studies with different age distributions.

In summary, while our results suggest that effects of established gene variants on lipid traits may vary with age, further studies are required to assess the validity of these findings. Due to the difficulties inherent in studying genetic modifiers, large collaborative efforts with longitudinal measures are necessary both to test for the existence of gene-environment interactions and to begin to understand the mechanisms behind them.

Acknowledgements

The PAGE program is funded by the National Human Genome Research Institute (NHGRI), supported by U01HG004803 (CALiCo), U01HG004798 (EAGLE), U01HG004802 (MEC), U01HG004790 (WHI), and U01HG004801 (Coordinating Center), and their respective NHGRI ARRA supplements. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The complete list of PAGE members can be found at http://www.pagestudy.org.

The data and materials included in this report result from a collaboration between the following studies: EAGLE is funded through the NHGRI PAGE program (U01HG004798–01 and its NHGRI ARRA supplement). Genotyping services for select NHANES III SNPs presented here were also provided by the Johns Hopkins University under federal contract number (N01-HV-48195) from National Heart, Lung, and Blood Institute (NHLBI). The study participants derive from the NHANES, and these studies are supported by the CDC. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the CDC.

The MEC study characterization of epidemiological architecture is funded through the NHGRI PAGE program (U01HG004802 and its NHGRI ARRA supplement). The MEC study is funded through the National Cancer Institute (R37CA54281, R01 CA63, P01CA33619, U01CA136792, and U01CA98758).

Funding support for the "Epidemiology of putative genetic variants: The Women's Health Initiative" study is provided through the NHGRI PAGE program (U01HG004790 and its NHGRI ARRA supplement).

The WHI program is funded by the NHLBI; NIH; and U.S. Department of Health and Human Services through contracts N01WH22110, 24152, 32100–2, 32105–6, 32108–9, 32111–13, 32115, 32118–32119, 32122, 42107–26, 42129–32, and 44221. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: https://cleo.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf.

Funding support for the Genetic Epidemiology of CALiCo program was provided through the NHGRI PAGE program (U01HG004803 and its NHGRI ARRA supplement). The following studies contributed to this manuscript and are funded by the following agencies: The ARIC Study is carried out as a collaborative study supported by NHLBI contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022. The CARDIA study is supported by the following National Institutes of Health, NHLBI contracts: N01-HC-95095; N01-HC-48047; N01-HC-48048; N01-HC-48049; N01-HC-48050; N01-HC-45134; N01-HC-05187; and N01-HC-45205. The CHS is supported by contracts HHSN268201200036C, 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 grant HL080295 from the NHLBI, with additional contribution from the 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.

Assistance with phenotype harmonization, SNP selection and annotation, data cleaning, data management, integration and dissemination, and general study coordination was provided by the PAGE Coordinating Center (U01HG004801–01 and its NHGRI ARRA supplement). The National Institutes of Mental Health also contributes to the support for the Coordinating Center.

The PAGE consortium thanks the staff and participants of all PAGE studies for their important contributions.

References

Anderson, G. L., Manson, J., Wallace, R., Lund, B., Hall, D., Davis, S., Shumaker, S., Wang, C. Y., Stein, E., & Prentice, R. L. (2003) Implementation of the women's health initiative study design. *Ann Epidemiol* **13**, S5–S17.

Arking, D. E., Pfeufer, A., Post, W., Kao, W. H. L., Newton-Cheh, C., Ikeda, M., West, K., Kashuk, C., Akyol, M., Perz, S., Jalilzadeh, S., Illig, T., Gieger, C., Guo, C. Y., Larson, M. G., Wichmann, H. E., Marban, E., O'Donnell, C. J., Hirschhorn, J. N., Kaab, S., Spooner, P. M., Meitinger, T., & Chakravarti, A. (2006) A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. *Nat Genet* **38**, 644–651.

Bonithon-Kopp, C., Scarabin, P. Y., Taquet, A., Touboul, P. J., Dame, B., & Guize, L. (1989) Increased risk of atherosclerosis in women after the menopause. *BMJ* **298**, 1311.

Boomsma, D. I., Kempen, H. J., Gevers Leuven, J. A., Havekes, L., de, K.P., & Frants, R. R. (1996) Genetic analysis of sex and generation differences in plasma lipid, lipoprotein, and apolipoprotein levels in adolescent twins and their parents. *Genet Epidemiol* **13**, 49–60.

Centers for Disease Control and Prevention (CDC) (2002) *National health and nutrition examination survey data*. Hyattsville, MD: U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention (2004) *Plan and operation of the third national health and nutrition examination survey, 1988–94*. Bethesda, MD: Department of Health and Human Services.

Centers for Disease Control and Prevention (2010) National health and nutrition examination survey (NHANES) DNA samples: Guidelines for proposals to use samples and cost schedule. *Fed Regist* **75**, 32191–32195.

Dumitrescu, L., Carty, C. L., Taylor, K., Schumacher, F. R., Hindorff, L. A., Ambite, J. L., Anderson, G., Best, L. G., Brown-Gentry, K., Buzkova, P., Carlson, C. S., Cochran, B., Cole, S. A., Devereux, R. B., Duggan, D., Eaton, C. B., Fornage, M., Franceschini, N., Haessler, J., Howard, B. V., Johnson, K. C., Laston, S., Kolonel, L. N., Lee, E. T., MacCluer, J. W., Manolio, T. A., Pendergrass, S. A., Quibrera, M., Shohet, R. V., Wilkens, L. R., Haiman, C. A., Le, M. L., Buyske, S., Kooperberg, C., North, K. E., & Crawford, D. C. (2011) Genetic determinants of lipid traits in diverse populations from the population architecture using genomics and epidemiology (PAGE) study. *PLoS Genet* **7**, e1002138.

Ericsson, S., Eriksson, M., Vitols, S., Einarsson, K., Berglund, L., & Angelin, B. (1991) Influence of age on the metabolism of plasma low density lipoproteins in healthy males. *J Clin Invest* **87**, 591–596.

Fried, L. P., Borhani, N. O., Enright, P., Furberg, C. D., Gardin, J. M., Kronmal, R. A., Kuller, L. H., Manolio, T. A., Mittelmark, M. B., Newman, A., O'Leary, D., Psaty, B., Rautaharju, P., & Tracy, R. (1991) The cardiovascular health study: Design and rationale. *Ann Epidemiol* **3**, 263–276.

Friedman, G. D., Cutter, G. R., Donahue, R. P., Hughes, G. H., Hulley, S. B., Jacobs, D. R., Liu, K., & Savage, P. J. (1988) CARDIA: Study design, recruitment and some characteristics of the examined subjects. *J Clin Epidemiol* **41**, 1105–1116.

Hardy, R. J. & Thompson, S. G. (1998) Detecting and describing heterogeneity in meta-analysis. *Stat Med* **17**, 841–856.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003) Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560.

Hindorff, L. A., MacArthur, J., Morales, J., Junkins, H. A., Hall, P. N., Klemm, A. K., & Manolio, T. A. (2011) A catalog of published genome-wide association studies. www.genome.gov/gwastudies. (accessed September 2012).

Hjortland, M. C., McNamara, P. M., & Kannel, W. B. (1976) Some atherogenic concomitants of menopause: The Framingham study. *Am J Epidemiol* **103**, 304–311.

Hunter, D. J. (2005) Gene-environment interactions in human diseases. *Nat Rev Genet* **6**, 287–298.

Jousilahti, P., Vartiainen, E., Tuomilehto, J., & Puska, P. (1996) Twenty-year dynamics of serum cholesterol levels in the middle-aged population of eastern Finland. *Ann Intern Med* **125**, 713–722.

Kolonel, L. N., Altshuler, D., & Henderson, B. E. (2004) The multiethnic cohort study: Exploring genes, lifestyle and cancer risk. *Nat Rev Cancer* **4**, 519–527.

Kronmal, R. A., Cain, K. C., Ye, Z., & Omenn, G. S. (1993) Total serum cholesterol levels and mortality risk as a function of age. A report based on the Framingham data. *Arch Intern Med* **153**, 1065–1073.

Mailman, M. D., Feolo, M., Jin, Y., Kimura, M., Tryka, K., Bagoutdinov, R., Hao, L., Kiang, A., Paschall, J., Phan, L., Popova, N., Pretel, S., Ziyabari, L., Lee, M., Shao, Y., Wang, Z. Y., Sirotkin, K., Ward, M., Kholodov, M., Zbicz, K., Beck, J., Kimelman, M., Shevelev, S., Preuss, D., Yaschenko, E., Graeff, A., Ostell, J., & Sherry, S. T. (2007) The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* **39**, 1181–1186.

Manolio, T. A., Bailey-Wilson, J. E., & Collins, F. S. (2006) Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* **7**, 812–820.

Matise, T. C., Ambite, J. L., Buyske, S., Carlson, C. S., Cole, S. A., Crawford, D. C., Haiman, C. A., Heiss, G., Kooperberg, C., Marchand, L. L., Manolio, T. A., North, K. E., Peters, U., Ritchie, M. D., Hindorff, L. A., & Haines, J. L. (2011) The next PAGE in understanding complex traits: Design for the analysis of population architecture using genetics and epidemiology (PAGE) study. *Am J Epidemiol* **174**, 849–859.

Matthews, K. A., Meilahn, E., Kuller, L. H., Kelsey, S. F., Caggiula, A. W., & Wing, R. R. (1989) Menopause and risk factors for coronary heart disease. *N Engl J Med* **321**, 641–646.

Ober, C. & Vercelli, D. (2011) Gene-environment interactions in human disease: Nuisance or opportunity? *Trends Genet* **27**, 107–115.

Reilly, S. L., Kottke, B. A., & Sing, C. F. (1990) The effects of generation and gender on the joint distributions of lipid and apolipoprotein phenotypes in the population at large. *J Clin Epidemiol* **43**, 921–940.

Shirts, B. H., Hasstedt, S. J., Hopkins, P. N., & Hunt, S. C. (2011) Evaluation of the gene-age interactions in HDL cholesterol, LDL cholesterol, and triglyceride levels: The impact of the SORT1 polymorphism on LDL cholesterol levels is age dependent. *Atherosclerosis* **217**, 139–141.

Snieder, H., van Doornen, L. J., & Boomsma, D. I. (1997) The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* **60**, 638–650.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P., Ripatti, S., Chasman, D. I., Willer, C. J., Johansen, C. T., Fouchier, S. W., Isaacs, A., Peloso, G. M., Barbalic, M., Ricketts, S. L., Bis, J. C., Aulchenko, Y. S., Thorleifsson, G., Feitosa, M. F., Chambers, J., Orho-Melander, M., Melander, O., Johnson, T., Li, X., Guo, X., Li, M., Shin, C. Y., Jin, G. M., Jin, K. Y., Lee, J. Y., Park, T., Kim, K., Sim, X., Twee-Hee, O. R., Croteau-Chonka, D. C., Lange, L. A., Smith, J. D., Song, K., Hua, Z. J., Yuan, X., Luan, J., Lamina, C., Ziegler, A., Zhang, W., Zee, R. Y., Wright, A. F., Witteman, J. C., Wilson, J. F., Willemsen, G., Wichmann, H. E., Whitfield, J. B., Waterworth, D. M., Wareham, N. J., Waeber, G., Vollenweider, P., Voight, B. F., Vitart, V., Uitterlinden, A. G., Uda, M., Tuomilehto, J., Thompson, J. R., Tanaka, T., Surakka, I., Stringham, H. M., Spector, T. D., Soranzo, N., Smit, J. H., Sinisalo, J., Silander, K., Sijbrands, E. J., Scuteri, A., Scott, J., Schlessinger, D., Sanna, S., Salomaa, V., Saharinen, J., Sabatti, C., Ruokonen, A., Rudan, I., Rose, L. M., Roberts, R., Rieder, M., Psaty, B. M., Pramstaller, P. P., Pichler, I., Perola, M., Penninx, B. W., Pedersen, N. L., Pattaro, C., Parker, A. N., Pare, G., Oostra, B. A., O'Donnell, C. J., Nieminen, M. S., Nickerson, D. A., Montgomery, G. W., Meitinger, T., McPherson, R., McCarthy, M. I., McArdle, W., Masson, D., Martin, N. G., Marroni, F., Mangino, M., Magnusson, P. K., Lucas, G., Luben, R., Loos, R. J., Lokki, M. L., Lettre, G., Langenberg, C., Launer, L. J., Lakatta, E. G., Laaksonen, R., Kyvik, K. O., Kronenberg, F., Konig, I. R., Khaw, K. T., Kaprio, J., Kaplan, L. M., Johansson, A., Jarvelin, M. R., Janssens, A. C., Ingelsson, E., Igl, W., Kees, H. G., Hottenga, J. J., Hofman, A., Hicks, A. A., Hengstenberg, C., Heid, I. M., Hayward, C., Havulinna, A. S., Hastie, N. D., Harris, T. B., Haritunians, T., Hall, A. S., Gyllensten, U., Guiducci, C., Groop, L. C., Gonzalez, E., Gieger, C., Freimer, N. B., Ferrucci, L., Erdmann, J., Elliott, P., Ejebe, K. G., Doring, A., Dominiczak, A. F., Demissie, S., Deloukas, P., de Geus, E. J., de, F.U., Crawford, G., Collins, F. S., Chen, Y. D., Caulfield, M. J., Campbell, H., Burtt, N. P., Bonnycastle, L. L., Boomsma, D. I., Boekholdt, S. M., Bergman, R. N., Barroso, I., Bandinelli, S., Ballantyne, C. M., Assimes, T. L., Quertermous, T., Altshuler, D., Seielstad, M., Wong, T. Y., Tai, E. S., Feranil, A. B., Kuzawa, C. W., Adair, L. S., Taylor, H. A., Jr., Borecki, I. B., Gabriel, S. B., Wilson, J. G., Holm, H., Thorsteinsdottir, U., Gudnason, V., Krauss, R. M., Mohlke, K. L., Ordovas, J. M., Munroe, P. B., Kooner, J. S., Tall, A. R., Hegele, R. A., Kastelein, J. J., Schadt, E. E., Rotter, J. I., Boerwinkle, E., Strachan, D. P., Mooser, V., Stefansson, K., Reilly, M. P., Samani, N. J., Schunkert, H., Cupples, L. A., Sandhu, M. S., Ridker, P. M., Rader, D. J., van Duijn, C. M., Peltonen, L., Abecasis, G. R., Boehnke, M., & Kathiresan, S. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713.

The ARIC Investigators (1989) The atherosclerosis risk in communities (ARIC) study: Design and objectives. *Am J Epidemiol* **129**, 687–702.

The Women's Health Initiative Study Group (1998) Design of the women's health initiative clinical trial and observational study. *Control Clin Trials* **19**, 61–109.

Third Report of the National Cholesterol Education Program (NCEP) Expert Panel (2002) Third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report. *Circulation* **106**, 3143–3421.

Webster, R. J., Warrington, N. M., Weedon, M. N., Hattersley, A. T., McCaskie, P. A., Beilby, J. P., Palmer, L. J., & Frayling, T. M. (2009) The association of common genetic variants in the APOA5, LPL and GCK genes with longitudinal changes in metabolic and cardiovascular traits. *Diabetologia* **52**, 106–114.

Willer, C. J., Li, Y., & Abecasis, G. R. (2010) METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 List of candidate gene and GWAS-identified SNPs genotyped in PAGE.

Table S2 Mean HDL-C levels across the six different PAGE studies.

Table S3 Mean LDL-C levels across the six different PAGE studies.

Table S4 Mean TG levels across the six different PAGE studies.

Table S5. Heterogeneity P-values for all analyses.

Received: 27 September 2012
Accepted: 28 March 2013