

Linkage Analysis of Plasma ApoE in Three Ethnic Groups: Multiple Genes with Context-Dependent Effects

K. L. E. Klos^{1,*}, S. L. R. Kardia², J. E. Hixson¹, S. T. Turner³, C. Hanis¹, E. Boerwinkle¹ and C. F. Sing⁴

¹Human Genetics Center, University of Texas Health Science Center, Houston, TX

²Department of Epidemiology, University of Michigan, Ann Arbor, MI

³Mayo Clinic, Rochester, MN

⁴Department of Human Genetics, University of Michigan, Ann Arbor, MI

Summary

We performed variance component-based linkage analysis in four samples (two of non-Hispanic European-Americans from Rochester, MN; African-Americans from Jackson, MS; and Mexican-Americans from Starr County, TX) to identify chromosomal regions containing genes influencing plasma apolipoprotein E (apoE) levels. The *APOE* gene region on chromosome (chr) 19 was identified with a $LOD \geq 2.00$ in both samples from Rochester and the sample from Jackson. Adjustment of apoE levels for differences among means of genotypes defined by the *APOE* $\epsilon 2/3/4$ alleles reduced evidence of linkage, indicating that the *APOE* gene was responsible for the majority of the linkage signal. In stratified linkage analyses, there was a LOD of 1.70 in the Starr County sibships with average total cholesterol (TC) above the median level for all sibships in that population. Adjustment for *APOE* genotype did not remove this LOD score, suggesting a second gene in this region may influence apoE variation. Evidence of linkage ($LOD = 3.32$) on chr 17 was observed in the Starr County sibships with average TC below the median. Inter-individual variation in plasma apoE level may be influenced by variations in the structural gene, and at least one other gene whose effects differ among populations and are dependent on the influence of unmeasured genetic and environmental factors indexed by correlated measures of lipid metabolism.

Keywords: Genome scan, sample stratification, apolipoprotein E, *APOE*

Introduction

In blood, apolipoprotein E (apoE) plays a regulatory role in the receptor-mediated metabolism of risk factors for coronary artery disease, including triglyceride (TG)-rich chylomicrons, low-density lipoprotein (LDL), and high-density lipoprotein (HDL) particles (Davignon *et al.* 1988, 1993; Mahaney *et al.* 1995; Mahley *et al.* 2000). Our previously published variance components-based linkage analysis validated the influence of the *APOE* structural gene on inter-individual variation in plasma apoE levels, but failed to provide

significant evidence for additional genes (Klos *et al.* 2001). In the study presented here we have expanded genetic linkage analyses to identify genomic regions that influence plasma apoE level in three additional samples (an additional sample of non-Hispanic European-Americans from Rochester, MN; a sample of African-Americans from Jackson, MS; and a sample of Mexican-Americans from Starr County, TX). We found evidence of linkage in the region of the *APOE* structural gene (chromosome 19q) in all samples except Starr County, which was unique in that sibships were ascertained by a proband having type II diabetes. Tentative to suggestive evidence of linkage specific to a single sample was observed at six regions of the genome.

The biological reality of the complex multifactorial etiology of quantitative traits such as plasma apoE suggests that many genes may interact with each other, and

*Corresponding author: Dr. Kathy L.E. Klos University of Texas Health Science Center at Houston School of Public Health Human Genetics Center P.O. Box 20186 Houston, TX 77225 USA. Telephone: (713) 500-9914; Fax: (713) 500-0900. Email: kklos@sph.uth.tmc.edu

with environmental factors, resulting in a particular gene contributing to trait variation in some families but not others (Sing *et al.* 1996, 2003; Greenspan, 2001). The extent to which populations-of-inference are homogeneous for the relative frequencies of causal genetic and environmental agents will also influence the probability of replicating linkage evidence. This potential for context-dependency, along with the unique lack of evidence for linkage in the Starr County sample at the well-studied *APOE* gene region, led us to examine evidence of linkage within particular genetic and environmental contexts of this sample, as indexed by sibship average total cholesterol (TC) level. Individuals with type II diabetes have an altered lipid profile, implying that strata of pedigrees indexed by TC may have greater homogeneity of interacting genetic and environmental effects, resulting in stronger single gene-phenotype relationships. Stratified linkage analyses of the Starr County sample revealed evidence of context-dependency at the *APOE* gene region on chromosome (chr) 19. Analyses of all four samples stratified by TC revealed significant evidence of linkage on chr 17 in the low TC strata of Starr County, suggesting the presence of an additional gene or genes whose effects on plasma apoE levels were undetected in the unstratified analyses.

METHODS

Sample and Laboratory Methods

The four samples examined in this study came from two different collaborative efforts – the Rochester Family Heart Study (RFHS) and the Genetic Epidemiology Network of Arteriopathy (GENOA). As part of the RFHS, multi-generation pedigrees were ascertained without regard for health through households with two or more children enrolled in the primary and secondary schools of Rochester, MN. Sampling details, the clinic examination protocol and baseline characteristics have been described by Moll *et al.* (1989) and Turner *et al.* (1989). Genotype data for 372 highly polymorphic microsatellite marker loci located on the 22 human autosomes was obtained for 1484 individuals (779 females and 705 males) in 232 multi-generation pedigrees. Median distance between markers was 8.3 centiMorgans (cM). Genotyping methods and laboratory analyses are

described in detail elsewhere (Klos *et al.* 2001). Plasma apoE level was measured by radioimmunoassay (Kottke *et al.* 1991).

As part of the GENOA study, sibships from Jackson, MS and Rochester, MN were recruited if they contained at least two full siblings with essential hypertension, clinically diagnosed before the age of 60 years. Sibships from Starr County, TX were recruited if they contained at least 2 full siblings with type II diabetes. Sampling details, the clinic and laboratory protocols and baseline characteristics have been described by Daniels *et al.* (2004). Genotyping of 373 highly polymorphic microsatellite marker loci located on the 22 human autosomes (from CHLC/Weber screening set 9.0) was performed at the Mammalian Genotyping Center of the Marshfield Medical Research Foundation. Median distance between markers was 8.5 cM. Genotype and phenotype data were obtained for 1696 non-Hispanic African-Americans (1166 females and 530 males) in 592 sibships from Jackson (GENOA-Jackson), for 1415 non-Hispanic European-Americans (770 females and 645 males) in 505 sibships from Rochester (GENOA-Rochester), and for 1651 Mexican-American individuals (974 females and 677 males) in 418 sibships from Starr County (GENOA-Starr County). There were 369 half-sib and 1781 full-sib pairs in GENOA-Jackson, 50 half-sib and 1716 full-sib pairs in GENOA-Rochester, and 118 half-sib and 3022 full-sib pairs in GENOA-Starr County. Forty-six individuals (in 36 sibships) sampled as part of GENOA-Rochester were also sampled as part of the RFHS. Details of the marker genotyping and validation are reported elsewhere (Kardia *et al.* 2003). Plasma apoE level was measured by radioimmunoassay (Kottke *et al.* 1991).

Homogeneity of quantitative trait variances among samples was evaluated using Levene's test (Levene, 1960). Differences among sample means were evaluated by the one-way analysis of variance (Sokal & Rohlf, 1998). Where variances were unequal, Welch's modified F statistic (Welch, 1951) was used in assessing significance. Equivalent conclusions were obtained using the Welch technique and the standard analysis of variance. The Boerwinkle & Sing (1986) bias-corrected estimator of genetic variance was used to measure the percent of plasma apoE variance attributable to deviations of *APOE* genotype means from the population mean. An

alpha level of 0.05 was used to determine statistical significance, except where otherwise noted.

Linkage Analysis

Prior to linkage analysis, a linear regression model containing age, age², age³, height, height², height³, weight, weight², weight³, and BMI was fitted to plasma apoE level within gender and, after adjustment, the sample grand mean was added to the residuals. The methods used for linkage analysis are described in detail in Klos *et al.* (2001). In brief, using the computer program SOLAR (Almasy & Blangero, 1998), we fitted a linear mixed model to the age, height, weight and BMI adjusted data such that for each pedigree the phenotypic variance-covariance matrix (Ω) could be written:

$$\Omega = \Phi\sigma_{QTL}^2 + \Pi\sigma_R^2 + \mathbf{I}\sigma_E^2$$

where Φ is a matrix of the proportion of alleles shared identical by descent (IBD) at a point in the genome, σ_{QTL}^2 represents the portion of the phenotypic variance about the trait mean attributable to the contribution of a quantitative trait locus (QTL), Π is a matrix of the expected proportion of alleles shared IBD for pairs of relatives, σ_R^2 is the variance attributable to familial genetic and unspecified shared environmental effects, \mathbf{I} is an identity matrix, and σ_E^2 is the variance attributable to measurement errors and environmental effects unique to the individual. Due to the existence of large pedigrees in the RFHS (pedigree size ranged from 2 to 31 individuals), and to maintain consistency of analysis across samples, multipoint estimates of IBD were obtained from SOLAR as a weighted average of the IBD at each individual marker. Linkage analyses based on the IBD approximation method employed by SOLAR have lower power than those based on multipoint IBDs computed by a Markov chain Monte Carlo approach (Goring *et al.* 2002). However, the SOLAR method is more robust to genotyping error and to double recombination events. LOD scores were calculated as the difference between the maximum of the log₁₀ likelihoods of the full model, above, and of the reduced model in which σ_{QTL}^2 was constrained to equal zero (Fulker *et al.* 1995; Almasy & Blangero, 1998). Likelihood calculations were performed under the assumption of an underlying multivariate t-distribution that allows for robust estimates of

variance components in the presence of deviations from a normal distribution (Hopper, 1988; Lange *et al.* 1988, 1989).

In the presence of genetic heterogeneity, sample stratification may improve the power to detect linkage provided there is reduced heterogeneity within strata compared to heterogeneity among strata (Leal & Ott, 2000). However, identification of an appropriate stratification variable is problematic. We hypothesize that the genetic influences on plasma apoE variation may differ among groups defined by different plasma lipid profiles. Because a biologically relevant measure of the difference in the genetic and environmental influences on plasma apoE levels between dyslipidaemic and normolipidemic individuals is not known, sample stratification in this case was based on the pedigree average TC level. The average TC level of all members was calculated for each pedigree; pedigrees were ranked by average TC, and then divided at the median into two strata.

We have used a LOD score threshold of ≥ 3.00 to indicate statistically significant evidence of linkage, a LOD of ≥ 2.00 and < 3.00 as suggestive, and a LOD ≥ 1.30 and < 2.00 as tentative evidence of linkage (Lander & Kruglyak, 1995; Morton, 1998).

In order to test the hypothesis that the LOD score peaks on chr 19 were due to the influence of the *APOE* structural gene, the *APOE* $\epsilon 2/3/4$ polymorphism was genotyped in the individuals from all four samples, and plasma ApoE levels were adjusted for *APOE* genotype means (after previous adjustments for age, height, weight and BMI). This method removes plasma apoE variation associated with the fixed, marginal effects of *APOE* $\epsilon 2/3/4$ genotypes. Linkage analysis was performed on plasma ApoE levels adjusted for *APOE* genotype as described above. This method should remove evidence of linkage on chr 19 due to variation in the *APOE* $\epsilon 2/3/4$ polymorphism, and may increase evidence of linkage at other QTLs by increasing the proportion of variance that they explain. In a second method, the *APOE* $\epsilon 2/3/4$ polymorphism was added to the marker list at 70 cM on chr 19. Multipoint IBD estimates were obtained, incorporating the *APOE* $\epsilon 2/3/4$ polymorphism information, and variance component-based linkage analysis was performed on plasma apoE adjusted for age, height, weight and BMI. In this method, the proportion of variance attributed to IBD sharing for the *APOE*

Table 1 Characteristics of 1484 individuals from the Rochester Family Heart Study (RFHS), of 1696 individuals from the Genetic Epidemiology Network of Arteriopathy (GENOA) - Jackson, of 1415 individuals from GENOA-Rochester and of 1650 individuals from GENOA-Starr County

	RFHS	GENOA-Jackson	GENOA-Rochester	GENOA-Starr County
N	1484	1696	1415	1651
% Female	52.5	68.8	54.4	59.0
Pedigrees	232	592	505	418
Age	37.7±24.1 [†]	57±10.3	55.9±11.0	55.2±12.1
Height (cm)	163.8±14.8	168.7±8.8	168.7±9.3	162.9±9.2
Weight (kg)	64.7±20.2	87.8±18.7	86.7±19.9	81.8±17.3
Systolic BP	114.0±22.3	136.4±22.8	133.9±17.1	128.5±19.9
Diastolic BP	63.2±13.3	77.8±12.4	78.6±9.6	73.8±10.1
BMI (kg/m ²)	23.6±5.3	30.9±6.6	30.4±6.3	30.8±6.1
TC (mg/dl)	178.5±45.9	205.1±46.9	210.4±38.9	205.4±46.4
ApoE (mg/dl)	5.1±.9	5.3±2.2	5.2±1.7	5.4±2.2

*Significant differences among samples at an $\alpha = 0.05$

[†]Mean ± the standard deviation

$\epsilon 2/3/4$ polymorphism is estimated, and should result in increased power to detect the *APOE* gene effect.

Results

Descriptive Statistics

Characteristics of the 1484 RFHS, the 1696 GENOA-Jackson, the 1415 GENOA-Rochester and the 1651 GENOA-Starr County individuals used in the linkage analyses are reported in Table 1. Means and variances differed significantly among samples at the $p < 0.001$ level for all traits, largely due to the difference in sampling strategy between the RFHS and GENOA studies. The multi-generation RFHS pedigrees were younger ($p < 0.0001$), less obese ($p < 0.0001$), and had lower TC ($p < 0.0001$), than the average of the three GENOA samples of sibships. The mean apoE level of the GENOA-Starr County sample was significantly higher than the mean of the other three samples pooled ($p = 0.0009$). Heritability of plasma ApoE was 0.57 ± 0.07 in GENOA-Jackson, 0.40 ± 0.07 in GENOA-Rochester, 0.39 ± 0.06 in GENOA-Starr County and 0.62 ± 0.05 in the RFHS. Heritability estimates were significant at an $\alpha = 0.0001$ in all cases.

Linkage Analysis

In Figure 1, the multipoint LOD scores from linkage analyses of all four samples are plotted against map position for each chromosome (in cM from the p-terminal). The highest multipoint LOD scores from analyses of

GENOA-Jackson, GENOA-Rochester and the RFHS were on chr 19 ($LOD = 4.69$ at 69 cM, $LOD = 2.32$ at 74 cM and $LOD = 4.20$ at 78 cM, respectively). The peak multipoint LOD score on chr 19 in GENOA-Starr County was 0.97 at 78 cM. In addition to significant evidence of linkage on chr 19, a multipoint LOD score suggestive of linkage ($LOD = 2.23$ at 97 cM) was observed on chr 3 in the GENOA-Jackson sample. Tentative evidence of linkage was observed in the GENOA-Jackson sample on chr 2 ($LOD = 1.65$ at 210 cM) and chr 8 ($LOD = 1.60$ at 67 cM), in the GENOA-Rochester sample on chr 13 ($LOD = 1.44$ at 80 cM), in the GENOA-Starr County sample on chr 10 ($LOD = 1.68$ at 67 cM) and in the RFHS on chr 12 ($LOD = 1.50$ at 53 cM).

The *APOE* $\epsilon 2/3/4$ polymorphism accounted for 11% ($p < 0.0001$), 9% ($p < 0.0001$), 3% ($p = 0.0002$) and 1% ($p = 0.0182$) of plasma apoE variance in males of RFHS, GENOA-Jackson, GENOA-Rochester and GENOA-Starr County, respectively. In females, this polymorphism accounted for 11% ($p < 0.0001$), 20% ($p < 0.0001$) and 17% ($p < 0.0001$) of apoE variance in RFHS, GENOA-Jackson and GENOA-Rochester, but was not a significant source of variation (0.4%, $p = 0.1170$) in GENOA-Starr County. The multipoint LOD scores on chr 19 from linkage analyses of the adjusted apoE levels in all four samples are plotted in Figure 2A. After adjustment for *APOE* genotype, the LOD scores decreased in all samples. The highest multipoint LOD scores were 1.72 at 71 cM in

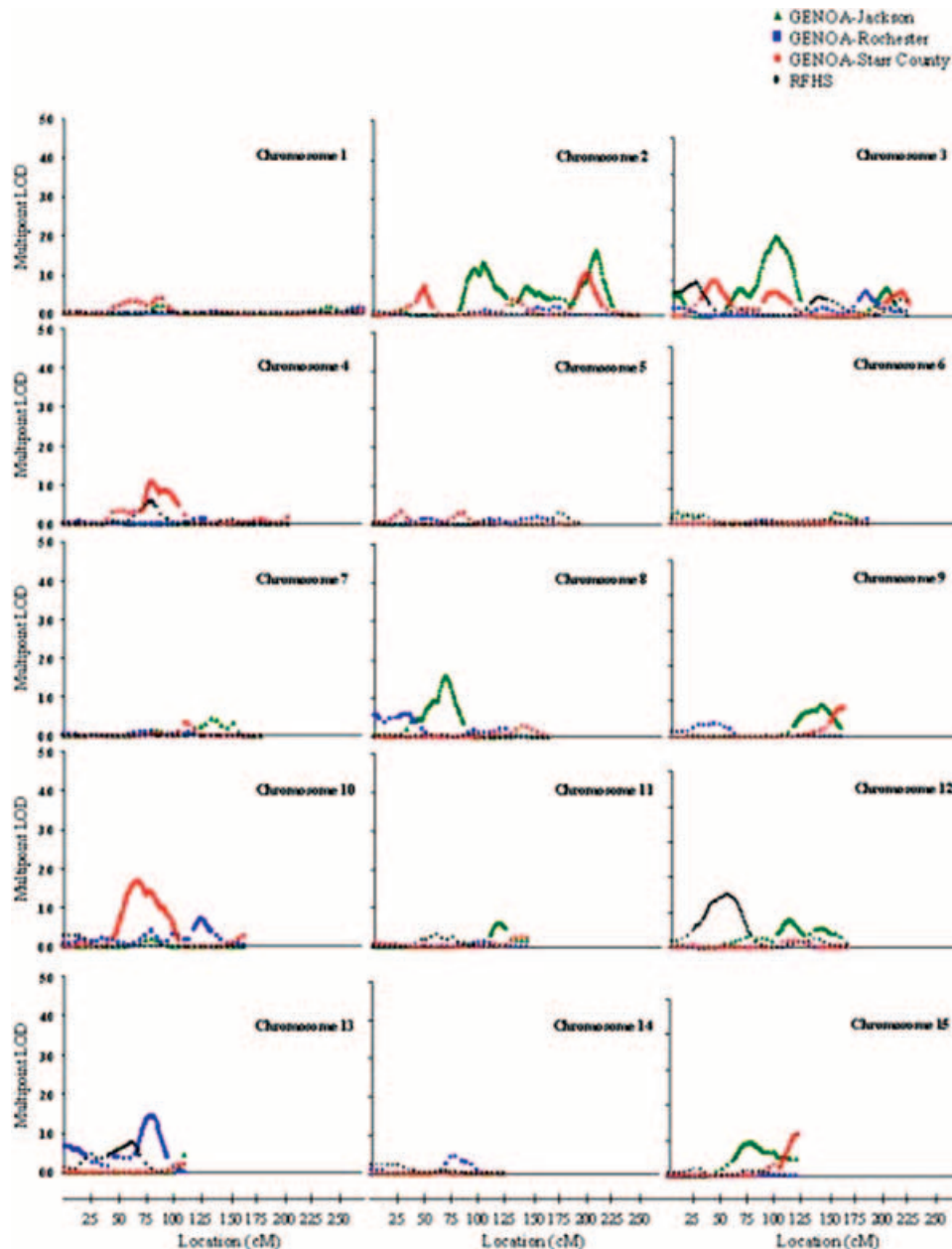


Figure 1 Multipoint LOD score plots for 22 chromosomes from the linkage analyses of plasma apoE level in pedigrees from the Rochester Family Heart Study (RFHS) and GENOA sibships from Jackson (GENOA-Jackson), Rochester (GENOA-Rochester) and Starr County (GENOA-Starr County).

GENOA-Jackson, 1.67 at 76 cM in GENOA-Rochester, and 0.80 at 78 cM in GENOA-Starr County (LOD scores for unadjusted data were 4.69 at 96 cM, 2.32 at 74 cM and 0.97 at 78 cM, respectively). This is consistent with previously documented evidence of a decrease in LOD score from 4.20 at 78 cM to 1.10 at 78 cM in the RFHS⁵, and indicates that in samples from Rochester and Jackson, variation in the *APOE*

gene accounted for the majority of the linkage signal on chr 19.

Evidence of linkage on chr 3 in GENOA-Jackson remained suggestive after adjustment of plasma apoE levels for *APOE* $\epsilon 2/3/4$ genotype means ($LOD = 2.20$ at 170 cM). Tentative evidence of linkage also remained on chr 2, 8, 10 and 13 (not shown). Tentative evidence of linkage on chr 12 in the RFHS did

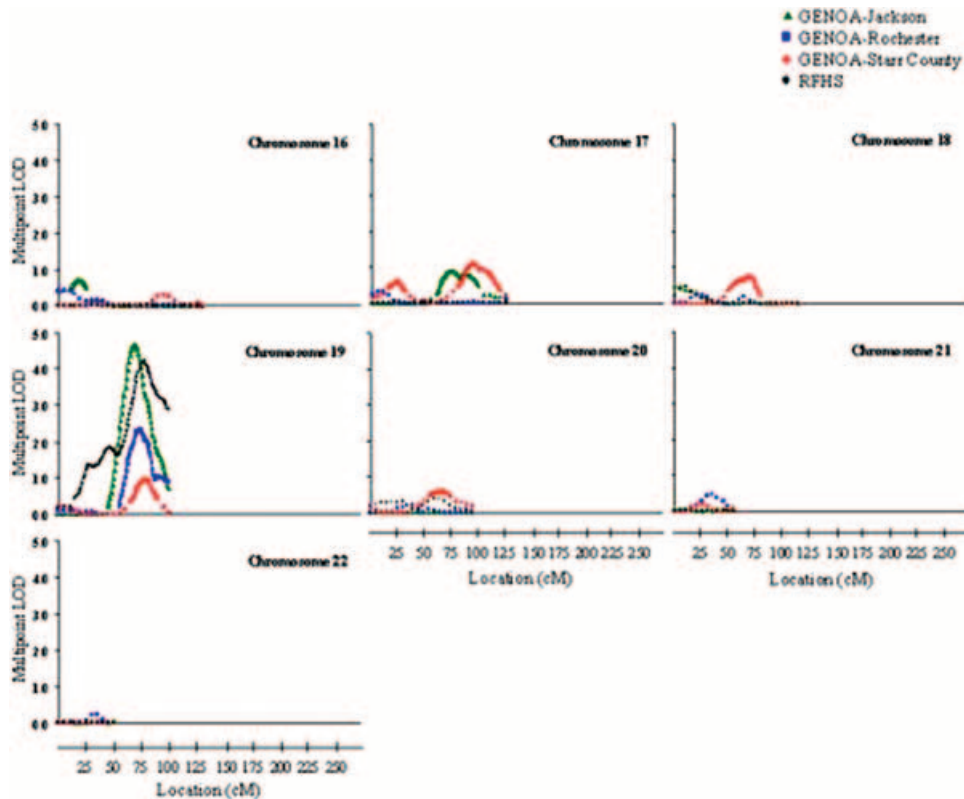


Figure 1 Continued.

not remain after adjustment for *APOE* genotype means ($LOD < 1.00$).

Linkage analyses with the measured *APOE* polymorphism incorporated into the marker set at 70 cM on chr 19 (using plasma apoE data unadjusted for *APOE* genotype) resulted in higher peak multipoint LOD scores in all samples except GENOA-Starr County. Inclusion of the *APOE* polymorphism resulted in the peak multipoint LOD scores on chr 19 increasing from 4.20 at 78 cM to 7.19 at 70 cM in the RFHS, from 4.69 at 96 cM to 6.09 at 70 cM in GENOA-Jackson, and from 2.32 at 74 cM to 2.51 at 73 cM in GENOA-Rochester. The peak multipoint LOD score remained 0.97 at 78 cM in GENOA-Starr County. The high LOD scores in GENOA-Jackson and the RFHS provide strong support for the hypothesis that the *APOE* structural gene is a major determinant of plasma apoE levels in these samples.

Stratified Linkage Analyses

The well-documented context-dependency of the phenotypic effects of the *APOE* gene (Zerba *et al.* 1996;

Jarvik *et al.* 1997; Kardia *et al.* 1998, 1999; Campos *et al.* 2001; Lussier-Cacan *et al.* 2002) supports the hypothesis that the control of plasma apoE level depends on multiple genetic and environmental factors. In GENOA-Starr County, the peak multipoint LOD score on chr 19 was below the threshold of tentative evidence of linkage, indicating that the *APOE* gene may play little or no role in determining inter-individual variation in plasma apoE levels, in this particular population, or may influence plasma apoE levels in some pedigrees but not others. To evaluate the potential of context-dependent stratified linkage analysis for detecting heterogeneity of the effects of the *APOE* gene, total cholesterol was selected as an index of the relative level of dyslipidaemia among the sibships of GENOA-Starr County. On chr 19, a LOD score of 1.70 at 83 cM (tentative evidence of linkage) was observed for GENOA-Starr County sibships with average TC levels above the median (figure 2B). Evaluation of linkage in these sibships incorporating the *APOE* $\epsilon 2/3/4$ polymorphism as part of the marker map did not improve LOD scores (peak LOD on chr 19 = 1.68 at 84 cM). Stratified linkage analyses on chr 19 using

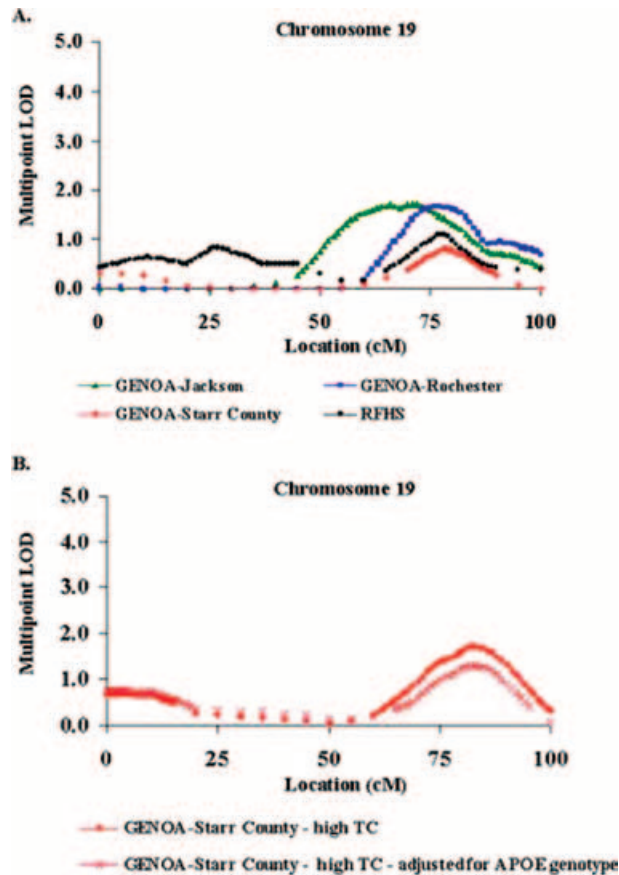


Figure 2 Multipoint LOD scores from linkage analyses of plasma apoE level adjusted for *APOE* $\epsilon 2/3/4$ genotype on chromosome 19 A) in pedigrees from the Rochester Family Heart Study (RFHS) and GENOA sibships from Jackson (GENOA-Jackson), Rochester (GENOA-Rochester) and Starr County (GENOA-Starr County) and B) in sibships from GENOA-Starr County with average TC level above the median.

the plasma apoE levels adjusted for age, height, weight, BMI and *APOE* genotype only reduced the peak LOD score to 1.30 at 83 cM, still tentative evidence of linkage (figure 2B).

Linkage analyses of all chromosomes were performed in all samples stratified at the median by sibship or pedigree average TC level. Peak multipoint LOD scores ≥ 1.30 from the stratified linkage analyses are shown in Table 2. Statistically significant evidence of linkage on chr 17 ($LOD = 3.32$ at 103 cM) was observed in GENOA-Starr County sibships with TC levels below the median (Figure 3). In the other three samples, peak multipoint LOD scores on chr 17 in strata indexed by pedigree average TC were below the threshold of tentative evidence of linkage (not shown), as were those

of the unstratified analyses of this region (Figure 1). In contrast to GENOA-Starr County, LOD scores on chr 19 from the stratified analyses of the GENOA-Jackson, GENOA-Rochester and RFHS samples were consistently lower than those in the unstratified analyses.

Discussion

Evidence of Linkage in the *APOE* Gene Region

Our linkage analyses of the *APOE* gene region document the dependency of the *APOE* gene on the background genetic and environmental context that varies among populations, and among pedigrees within a particular population. Replicated evidence of the influence of the *APOE* gene in three out of four samples is consistent with the many studies of the impact of this gene on lipid metabolism. As much as 50% of genotypic variation in plasma apoE level may be attributable to variation in the *APOE* gene (Boerwinkle & Utermann, 1988; Neale *et al.* 2000). Adjusting plasma apoE for the observed *APOE* genotypes influenced the evidence of linkage on chr 19 in all samples. In GENOA-Jackson and GENOA-Rochester, tentative evidence of linkage remained after adjustment for *APOE* genotype ($LOD = 1.72$ and 1.67 , respectively). Adjustment of plasma apoE levels was made using only the $\epsilon 2/3/4$ polymorphism, and did not account for the effects of other functional polymorphisms in the *APOE* gene region. The *APOE* gene is part of a 48 kb cluster including the *APOC1*, *APOC4* and *APOC2* genes, and two hepatic control regions which influence expression of all genes in the cluster (Jong *et al.* 1999). Variation in the *APOE/C1/C2/C4* gene cluster, outside the *APOE* coding region, may also be partially responsible for inter-individual variation in plasma apoE level in these samples.

The effect of the *APOE* gene on plasma apoE level is not uniform among or within populations. Depending on the gender and population studied, the *APOE* gene region may explain as much as 30%, or as little as 5% of phenotypic variation in plasma apoE levels (Davignon *et al.* 1988; Kaprio *et al.* 1991; Hsueh *et al.* 2000; Sten-gard *et al.* 2002). The lack of evidence for the contribution of this gene region in the GENOA-Starr County sample may indicate an influence of the *APOE* gene on

Chromosome	LOD score	Location	Sample	Strata
2	1.42	179 cM	GENOA-Jackson	Low TC
3	1.42	213 cM	GENOA-Starr County	Low TC
6	1.51	0 cM	GENOA-Jackson	Low TC
6	1.44	30 cM	RFHS	High TC
7	1.90	137 cM	GENOA-Starr County	Low TC
12	1.31	36 cM	RFHS	High TC
15	1.64	122 cM	GENOA-Starr County	High TC
17	3.32	103 cM	GENOA-Starr County	Low TC
19	2.29	68 cM	GENOA-Jackson	Low TC
19	2.49	71 cM	GENOA-Jackson	High TC
19	3.20	71 cM	RFHS	Low TC
19	2.10	76 cM	GENOA-Rochester	Low TC
19	1.70	83 cM	GENOA-Starr County	High TC
19	1.99	88 cM	RFHS	High TC

Table 2 Peak multipoint LOD scores ≥ 1.30 for plasma apoE level from linkage analyses of the RFHS, GENOA-Jackson, GENOA-Rochester and GENOA-Starr County samples stratified at the median by pedigree average TC level.

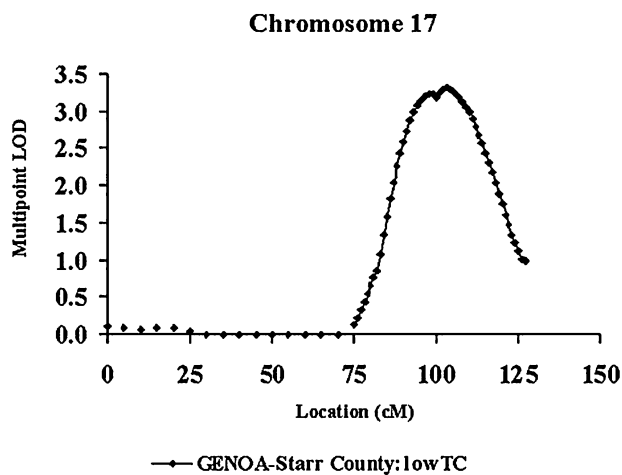


Figure 3 Multipoint LOD scores from the linkage analysis of plasma apoE on chromosome 17 in GENOA sibships from Starr County (GENOA-Starr County) with average total cholesterol (TC) levels below the median.

plasma apoE below the level detectable by linkage analysis. This seems likely in light of our subsequent stratified analyses.

Within-population context-dependencies of the phenotypic effect of the *APOE* gene on plasma apoE level have been documented, and include age, smoking and BMI (Zerba *et al.* 1996; Jarvik *et al.* 1997; Lussier-Cacan *et al.* 2002). Stratification can provide a means of incorporating context-dependency in such a way as to reveal evidence of linkage not obtained from the full sample. Leal & Ott (2000) used stratification to improve the ability to detect linkage in the presence of genetic heterogeneity among samples of simulated affected-sib-pairs, where variation at a locus contributed to phenotypic variation in some, but not all, sibships. Improved

LOD scores for diastolic and systolic blood pressure were found by Rice *et al.* (2000) when pedigrees ascertained through a highly obese (BMI > 32 kg/m²) family member were separated from randomly ascertained pedigrees. Stratified linkage analyses of the GENOA-Starr County sample indexed by sibship average TC level provided tentative evidence of linkage in the region of the *APOE* gene. However, adjustment of plasma apoE level for *APOE* genotype in GENOA-Starr County sibships with average TC above the median did not remove evidence of linkage, indicating that the $\epsilon 2/3/4$ polymorphism was not responsible for the linkage signal. The context-dependent linkage evidence on chr 19 in GENOA-Starr County may be due to type I error, or may indicate the context-dependent influence of genetic variation other than *APOE* $\epsilon 2/3/4$ on plasma apoE levels. Examples in yeast and maize provide precedents for multiple genes underlying a single QTL (Steinmetz *et al.* 2002; Zhang *et al.* 2003). In GENOA-Starr County, identification of new linkage evidence on chr 19, with stratification by TC, serves to emphasize the importance of context-dependent gene effects in defining the genetic architecture of a quantitative intermediate risk factor of CAD. Our linkage analyses of the chr 19 *APOE* gene region also highlight the difficulties to be faced as researchers attempt to identify the genes responsible for a linkage peak.

Evidence of Linkage on Other Chromosomal Regions

Suggestive evidence of linkage on chr 3 in the unstratified GENOA-Jackson sample may indicate a gene with

a weak or context-dependent effect on plasma apoE level. No obvious candidate genes lie in this region, but this peak is one cM from the location, identified by Imperatore *et al.* (2000), of tentative evidence of linkage with a gene(s) influencing TG levels (LOD = 1.77 at 98 cM) in a sample of Pima Indians. Tentative evidence of linkage in GENOA–Starr County sibships with average TC below the median (Table 2) was 116 cM from the GENOA–Jackson peak, and is not likely to represent the same underlying gene(s). Stratification of the GENOA–Jackson sample by sibship average TC level did not improve the LOD score in this region.

With stratification of the GENOA–Starr County sample by TC, there was significant evidence of linkage on chr 17 with a gene(s) that may influence plasma apoE levels in a manner dependent on genetic and environmental context, at least partially indexed by TC. A strong candidate, the *APOH* gene, is located near the multipoint LOD score peak on chr 17. Variation in the *APOH* gene has been associated with variation in plasma apoE levels in unrelated males from the RFHS (Kaprio *et al.* 1991) and may also influence TG, TC and LDL-C levels in a population-dependent manner (Kamboh *et al.* 1996, 1999). To our knowledge, within-population context-dependency of the *APOH* gene has not been investigated. Further study is warranted to investigate the influence of variation in *APOH*, or other genes in this region of chr 17, on variation in plasma apoE levels.

With the exception of chr 19 (Figure 1), and the possible exception of tentative evidence of linkage on chr 6 (Table 2), no linkage peak in these stratified or unstratified analyses of plasma apoE level had a LOD score ≥ 1.30 in a second sample. Replicated evidence of linkage has proven notoriously difficult to procure in the analysis of complex traits; however, this would be expected if gene effects are dependent on sample-specific (or strata-specific) genetic and environmental contexts. While evidence of linkage on both chr 3 and 17 is supported by evidence from additional linkage and candidate gene studies, type I error remains a possibility due to the lack of replication in independent analyses.

It has been argued that combined analysis of many samples may be necessary to provide the power to detect genes of small effect in complex traits (Ioannidis *et al.* 2001; Lohmueller *et al.* 2003; Province *et al.* 2003).

If complex traits are governed by context-dependent gene effects (which may be quite large in some pedigrees), rather than by small invariant effects, then a meta-analysis will only serve to exacerbate the problem by burying the signal under an even larger pile of non-informative pedigrees. This may explain the failure of some meta-analyses to provide support for previously identified evidence of linkage in studies of complex traits such as hypertension and asthma (Province *et al.* 2003; Etzel & Guerra, 2002). Our stratified linkage results on chr 17 and 19 support the expectation that identification of some genes which influence plasma apoE levels will require a statistical model which incorporates genetic and environmental context. We conclude that research strategies that reflect the biological reality of etiological complexity may provide more valuable insights into the genetic architecture of quantitative traits than studies that simply test for invariant, universal genetic effects.

Acknowledgments

We wish to thank Paul Kopec and Kenneth G. Weiss for their dedication to competently modifying and executing the database management and linkage analysis programs used in the research reported here. Genotyping and linkage analyses were supported by NIH/National Heart Lung Blood Institute grants HL-39107, HL-54457 and HL-51021. The collection of DNA samples and phenotype data were supported by NIH/National Heart Lung Blood Institute grants HL-54464, HL-54463, HL-54504, HL-54481 and HL-54457.

References

- Almasy, L. & Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* **62**, 1198–1211.
- Boerwinkle, E. & Sing, C. F. (1986) Bias of the contribution of single locus effects to the variance of a quantitative trait. *Am J Hum Genet* **39**, 137–144.
- Boerwinkle, E. & Utermann, G. (1988) Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet* **42**, 104–112.
- Campos, H., D’Agostino, M. & Ordovas, J. M. (2001) Gene-diet interactions and plasma lipoproteins: Role of apolipoprotein E and habitual saturated fat intake. *Genet Epidemiol* **20**, 117–128.
- Daniels, P. R., Kardia, S. L. R., Hanis, C. L., Brown, C. A., Hutchinson, R., Boerwinkle, E., Turner, S. T.; Genetic Epidemiology Network of Arteriopathy study. (2004) The

- familial aggregation of hypertension treatment and control in the GENOA study. *Am J Med* **116**, 676–81.
- Davignon, J., Gregg, R. E. & Sing, C. F. (1988) Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* **8**, 1–21.
- Davignon, J. (1993) Apolipoprotein E polymorphism and atherosclerosis. In: *New horizons in coronary heart disease*. London: Current Science.
- Etzel, C. J., Guerra, R. (2002) Meta-analysis of genetic-linkage analysis of quantitative-trait loci. *Am J Hum Genet* **71**, 56–65.
- Fulker, D. W., Cherny, S. S. & Cardon, L. R. (1995) Multi-point interval mapping of quantitative trait loci, using sib pairs. *Am J Hum Genet* **56**, 1224–1233.
- Görling, H. H., Williams, J. T., Dyer, T. D. & Blangero, J. (2002) On different approximations to multilocus identity-by-descent calculations and the resulting power of variance component-based linkage analysis. *BMC Genet* **4**(Suppl 1): S72.
- Greenspan, R. J. (2001) The flexible genome. *Nat Rev Genet* **2**, 383–387.
- Hopper, J. L. (1988) Review of FISHER. *Genet Epidemiol* **5**, 473–476.
- Hsueh, W., Mitchel, B. D., Hixson, J. E. & Rainwater, D. L. (2000) Effects of the ApoE polymorphism on plasma lipoproteins in Mexican Americans. *Ann Epidemiol* **10**, 524–531.
- Imperatore, G., Knowler, W. C., Pettitt, D. J., Kobes, S., Fuller, J. H., Bennett, P. H. & Hanson, R. L. (2000) A locus influencing total serum cholesterol on chromosome 19p. Results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arterioscler Thromb Vasc Biol* **20**, 2651–2656.
- Ioannidis, J. P. A., Ntzani, E. E., Trikalinos, T. A. & Contopoulos-Ioannidis, D. G. (2001) Replication validity of genetic association studies. *Nat Genet* **29**, 306–309.
- Jarvik, G. P., Goode, E. L., Austin, M. A., Auwerx, J., Deeb, S., Schellenberg, G. D. & Reed, T. (1997) Evidence that the apolipoprotein E - genotype effects on lipid levels can change with age in males: A longitudinal analysis. *Am J Hum Genet* **61**, 171–181.
- Jong, M. C., Hofker, M. H. & Havekes, L. M. (1999) Role of apoCs in lipoprotein metabolism. Functional differences between ApoC1, ApoC2, and ApoC3. *Arterioscler Thromb Vasc Biol* **19**, 472–484.
- Kamboh, M. I., Crawford, M. H., Aston, C. E. & Leonard, W. R. (1996) Population distributions of APOE, APOH, and APOA4 polymorphisms and their relationships with quantitative plasma lipid levels among the Evenki herders of Siberia. *Hum Biol* **68**, 231–243.
- Kamboh, M. I., Bunker, C. H., Aston, C. E., Nestlerode, C. S., McAllister, A. E. & Ukoli, F. A. (1999) Genetic association of five apolipoprotein polymorphisms with serum lipoprotein-lipid levels in African blacks. *Genet Epidemiol* **16**, 205–222.
- Kaprio, J., Ferrell, R. E., Kottke, B. A., Kamboh, M. I. & Sing, C. F. (1991) Effects of polymorphisms in apolipoproteins E, A-IV, and H on quantitative traits related to risk for cardiovascular disease. *Arterioscler Thromb* **11**, 1330–1348.
- Kardia, S. L., Haviland, M. B. & Sing, C. F. (1998) Correlates of Family History of Coronary Artery Disease in Children. *J Clin Epidemiol* **51**, 473–486.
- Kardia, S. L. R., Haviland, M. B., Ferrell, R. E. & Sing, C. F. (1999) The relationship between risk factor levels and presence of coronary artery calcification is dependent on *Apolipoprotein E* genotype. *Arterioscler Thromb Vasc Biol* **19**, 427–435.
- Kardia, S. L. R., Rozek, L. S., Krushkal, J., Ferrell, R. E., Turner, S. T., Hutchinson, R., Brown, A., Sing, C. F. & Boerwinkle, E. (2003) Genome-wide linkage analyses for hypertension genes in two ethnically and geographically diverse populations. *Am J Hypertens* **16**, 154–157.
- Klos, K. L. E., Kardia, S. L. R., Ferrell, R. E., Turner, S. T., Boerwinkle, E. & Sing, C. F. (2001) Genome-wide linkage analysis reveals evidence of multiple regions that influence variation in plasma lipid and apolipoprotein levels associated with risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* **21**, 971–978.
- Kottke, B. A., Moll, P. P., Michels, V. V. & Weidman, W. H. (1991) Levels of lipids, lipoproteins, and apolipoproteins in a defined population. *Mayo Clin Proc* **66**, 1198–1208.
- Lander, E. & Kruglyak, L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* **11**, 241–247.
- Lange, K., Weeks, D. & Boehnke, M. (1988) Programs for pedigree analysis: MENDEL, FISHER and dGENE. *Genet Epidemiol* **5**, 471–472.
- Lange, K. L., Little, R. J. A. & Taylor, J. M. G. (1989) Robust statistical modeling using the t-distribution. *JASA* **84**, 881–896.
- Leal, S. M. & Ott, J. (2000) Effects of stratification in the analysis of affected-sib-pair data: Benefits and costs. *Am J Hum Genet* **66**, 567–575.
- Levene, H. (1960) Robust tests for equality of variance. In: *Contributions to probability and statistics*. Stanford, CA: Stanford University Press.
- Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. H. (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* **33**, 177–182.
- Lussier-Cacan, S., Bolduc, A., Xhignesse, M., Niyonsenga, T. & Sing, C. F. (2002) Impact of alcohol intake on measures of lipid metabolism depends on context defined by gender, body mass index, cigarette smoking, and

- apolipoprotein E genotype. *Arterioscler Thromb Vasc Biol* **22**, 824–831.
- Mahaney, M. C., Blangero, J., Rainwater, D. L., Comuzzie, A. G., VandeBerg, J. L., Stern, M. P., MacCluer, J. W. & Hixson, J. E. (1995) A major locus influencing plasma high-density lipoprotein cholesterol levels in the San Antonio Family Heart Study. *Arterioscler Thromb Vasc Biol* **15**, 1730–1739.
- Mahley, R. W. & Rall, S. C. (2000) Apolipoprotein E: Far More Than a Lipid Transport Protein. *Annu Rev Genomics Hum Genet* **01**, 507–537.
- Moll, P. P., Michels, V. V., Weidman, W. H. & Kottke, B. A. (1989) Genetic determination of plasma apolipoprotein AI in a population-based sample. *Am J Hum Genet* **44**, 124–139.
- Morton, N. E. (1998) Significance levels in complex inheritance. *Am J Hum Genet* **62**, 690–697.
- Neale, M. C., de Knijff, P., Havekes, L. M. & Boomsma, D. I. (2000) ApoE polymorphism accounts for only part of the genetic variation in quantitative apoE levels. *Genet Epidemiol* **18**, 331–340.
- Province, M. A., Kardia, S. L. R., Ranade, K., Rao, D. C., Thiel, B. A., Cooper, R. S., Risch, N., Turner, S. T., Cox, D. R., Hunt, S. C., Weder, A. B. & Boerwinkle, E. (2003) A meta-analysis of genome-wide linkage scans for hypertension: The National Heart, Lung and Blood Institute Family Blood Pressure Program. *Am J Hypertens* **16**, 144–147.
- Rice, T., Rankinen, T., Province, M. A., Chagnon, Y. C., Perusse, L., Borecki, I. B., Bouchard, C. & Rao, D. C. (2000) Genome-wide linkage analysis of systolic and diastolic blood pressure. The Quebec Family Study. *Circulation* **102**, 1956–1963.
- Sing, C. F., Haviland, M. B. & Reilly, S. L. (1996) Genetic architecture of common multifactorial diseases. In: *Ciba Foundation Symposium 197: Variation in the Human Genome*. Chichester, England: Wiley.
- Sing, C. F., Stengard, J. H. & Kardia, S. L. R. (2003) Genes, environment and cardiovascular disease. *Arterioscler Thromb Vasc Biol* **23**, 1190–1196.
- Sokal, R. R. & Rohlf, F. J. (1998) *Biometry: the principles and practice of statistics in biological research*. New York, NY: Freeman and Company.
- Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H. & Davis, R. W. (2002) Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330.
- Stengård, J. H., Clark, A. G., Weiss, K. M., Kardia, S., Nickerson, D. A., Salomaa, V., Ehnholm, C., Boerwinkle, E. & Sing, C. F. (2002) Contributions of 18 additional DNA sequence variations in the gene for apolipoprotein E to explaining variation in quantitative measures of lipid metabolism. *Am J Hum Genet* **71**, 501–517.
- Turner, S., Weidman, W., Michels, V., Reed, T., Ormson, C., Fuller, T. & Sing, C. F. (1989) Distribution of sodium-lithium countertransport and blood pressure in caucasians five to eighty-nine years of age. *Hypertension* **13**, 378–391.
- Welch, B. L. (1951) On the comparison of several mean values: an alternative approach. *Biometrika* **38**, 330–336.
- Williams, J. T. & Blangero, J. (1999) Power of variance component linkage analysis to detect quantitative trait loci. *Ann Hum Genet* **63**(6): 545–563.
- Zhang, P., Wang, Y., Zhang, J., Maddock, S., Snook, M. & Peterson, T. (2003) A maize QTL for silk maysin levels contains duplicated Myb-homologous genes which jointly regulate flavone biosynthesis. *Plant Mol Biol* **52**, 1–15.
- Zerba, K. E., Ferrell, R. E. & Sing, C. F. (1996) Genotype-environment interaction: Apolipoprotein E (*ApoE*) gene effects and age as an index of time and spatial context in the human. *Genet* **143**, 463–478.

Received: 18 July 2003

Accepted: 20 August 2004