Am. J. Hum. Genet. 43:311–321, 1988

Sources of Interindividual Variation in the Quantitative Levels of Apolipoprotein B in Pedigrees Ascertained through a Lipid Clinic

Gail Pairitz,* Jean Davignon,† Helene Mailloux,† and Charles F. Sing‡

*Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia; †Department of Lipid Metabolism and Atherosclerosis Research, Clinical Research Institute of Montreal, Quebec; and ‡Department of Human Genetics, University of Michigan, Ann Arbor

Summary

The quantitative level of apolipoprotein (apo) B associated with low-density lipoprotein (LDL) varies among individuals within the population. This variation in level of the LDL receptor ligand appears to have predictive value, and may have an etiologic role, in coronary artery disease. Complex segregation analysis was used to compare eight different models of transmission. This study confirms the existence of allelic variations at a single genetic locus with large effects on the interindividual variation in the level of the serum apo B associated with LDL. This is the first study to consider the possible effects of inherited polymorphic variation in the apo E molecule when analyzing the components of variation in apo B associated with LDL. Our analyses suggest that the common alleles coding for the apo E polymorphism act independently of the unmeasured single-gene locus characterized by this study.

Introduction

Apolipoprotein B (apo B) associated with the low-density lipoprotein molecule (LDL-B) is the ligand for the binding of low-density lipoprotein (LDL) to its receptor. This property gives LDL-B a critical role in the delivery of cholesterol to the peripheral tissues. A number of studies have reported an association between elevated levels of apo B or LDL-B and coronary artery disease (CAD). Avogaro et al. found that total plasma apo B levels were increased in male patients with CAD as compared with controls (Avogaro et al. 1978) and that apo B was a better discriminator of CAD than were lipid levels in older men (Avogaro et al. 1979). Riesen et al. (1980) and Schmidt et al. (1985) both found significantly increased apo B levels in males with CAD. In a sample of both males and females, Sniderman et al. (1980) found LDL-B to be a better discriminator between patients with positive and negative coronary angiograms than were plasma cholesterol, triglyceride, and LDL-cholesterol (LDL-C) in subjects with LDL-C <200 mg/dl. Whayne et al. (1981) found LDL-B to be a better discriminator between individuals with positive and negative angiograms in males with total cholesterol <265 mg/dl than were age, cholesterol, triglyceride, very-low-density lipoprotein cholesterol (VLDL-C), LDL-C, or high-density lipoprotein cholesterol (HDL-C). Of patients with high VLDL-C (type IV hyperlipoproteinemia), those who had levels of LDL-B >129 mg/dl had higher rates of atherosclerotic disease than those who had LDL-B levels <84 mg/dl (Sniderman et al. 1982).

Complex segregation analyses to determine the genetic model for interindividual variability in the quantitative level of apo B have been reported by three groups: Beaty et al. (1986), Hasstedt et al. (1987), and Amos et al. (1987). In a single Amish pedigree of 240 individuals, ascertained through a proband with sitosterolemia, Beaty et al. (1986) found statistically significant evidence for polygenic effects on LDL-B but did not find statistically significant evidence for the segregation of a single locus with a large effect. In contrast, in a study of 331 individuals from 36 pedigrees ascertained through individuals having a variety of traits associated with the manifestation of cardiovascular disease, Hasstedt et al. (1986) found evidence for the segregation of a single locus with a large effect on total apo B levels. This locus explained 43% of the interindividual vari-

Received August 17, 1987; revision received April 4, 1988.

Address for correspondence and reprints: Dr. Charles F. Sing, Department of Human Genetics, 4708 Medical Sciences II, University of Michigan Medical School, Ann Arbor, MI 48109.

0002-9297/88/4303-0012$02.00

ance in apo B levels adjusted for variation in age and sex. The relative frequency of the common allele, which was associated with lower apo B levels, was estimated as .847. The polygenic contribution, estimated to be 15% of the variation in apo B among individuals with the same genotype at the single locus having a large effect, was not significantly different from zero at the .05 level of probability. Amos et al. (1987) applied methods of major gene analysis to data collected in a single pedigree of 196 individuals ascertained through an adolescent girl with a heart murmur. Fifteen members had a history of heart disease, and six had experienced at least one myocardial infarction. Their results suggest that in this pedigree >40% of the variability of total serum apo B levels can be explained by the presence of two common alleles that have relatively small effects. Although these three studies clearly establish a role of genetic factors, they do not establish the generality of the contribution of a single gene with a major effect on apo B levels. Hence, further investigation of the causes of interindividual variation in apo B in additional samples will extend our understanding of the role of a single genetic factor. Further, none of these studies have assessed the role of the common genetic polymorphism at the locus coding for the E apolipoprotein, which is known to be associated with differences in the mean level of LDL-B (Sing and Davignon 1985; Utermann 1985). A genetic analysis of LDL-B after adjustment for this known source of genetic variation in LDL-B levels, as is done here, has not been previously reported.

We have used complex segregation analysis to evaluate the role of genetic factors in the determination of variability in LDL-B levels among 396 individuals from 83 pedigrees ascertained through probands identified by a lipid clinic. Separate analyses of LDL-B levels adjusted for variation due to age and sex (B;as), LDL-B levels adjusted for variation due to age, sex, and apo E genotype (B;ase), and LDL-B levels adjusted for variation due to age, sex, apo E genotype, and LDL-C (B;asec) are presented and compared. The objectives were to assess the separate effects of the apo E genetic polymorphism on LDL-B levels, to estimate the fraction of genetic variation in LDL-B level which is independent of genetic variation in apo E genotype, and to investigate the independence of the genetic control of LDL-C and LDL-B in families ascertained through probands attending a lipid clinic. We also examined whether there was evidence for heterogeneity of the effects of apo E genotypes, or allelic variation at the single locus detected by complex segregation analysis, on LDL-B levels among groups of families identified by patients with different lipid disorders.

Methods

Definition of Sample Studied

The pedigrees considered here were ascertained through a proband identified by the Lipid Clinic of the Clinical Research Institute of Montreal. Patients with hyperlipidemia, xanthomas and other manifestations of dyslipoproteinemias, premature coronary or peripheral atherosclerosis, as well as individuals with a family history of CAD, are referred to this clinic. Patients are encouraged to have their family members also examined at the clinic. All pedigrees consisting of two or more individuals identified through a proband and who were seen by the clinic were considered in this study. Individuals in these pedigrees with familial hypercholesterolemia (FH) and who were diabetic, hypothyroid, pregnant, using birth control pills or a lipid-lowering medication, or following a prescribed lipid-lowering diet were removed from the sample. The diagnosis of FH required an elevation of LDL above 250 mg/dl, the presence of tendon xanthomas in the individual or a first-degree relative, and a family history of CAD before 55 years of age. After removal of these individuals there remained 394 individuals distributed in 83 pedigrees. Of these individuals, 224 were classified as normolipidemic, 120 as type IV hyperlipoproteinemia (HLP; defined below), 13 as type IIa HLP, 11 as type IIb HLP, 18 as type III HLP, and three as type V HLP; five were unclassified mild hypertriglyceridemics. The classification of HLP was made at the Lipid Clinic, based on lipid profile, age, sex, and clinical criteria. Asymptomatic individuals who had levels of total cholesterol <240, triglycerides <150, VLDL-C <35, and LDL-C <190 were considered normolipidemic. Hyperlipidemic individuals were considered to be type IIa, IIb, III, IV, and V HLP based on which fraction of their lipoprotein was greater than the 95th percentile for age and sex (Beaumont et al. 1970; Heiss et al. 1980; American Heart Association 1984). The fractions increased are LDL in type IIa, LDL and VLDL in type IIb, detectable beta-VLDL in type III, VLDL in type IV, and both chylomicrons and VLDL in type V. Five of the hyperlipidemic individuals had normal lipoprotein profiles but mild hypertriglyceridemia. To avoid misclassification, they were not included in analyses involving comparisons among diagnostic categories. There were 201 males and 193 females. Fifty-four individuals were ages

2–17 years, 118 were 18–29, 71 were 30–39, 71 were 40–49, 45 were 50–59, 20 were 60–69, and 15 were 70–84. Twelve of the pedigrees contained individuals from 3 generations, 53 pedigrees contained 2 generations, and 18 pedigrees contained 1 generation. Twenty-nine of the pedigrees contained 1–3 individuals, 42 contained 4–6 individuals, 11 contained 7–12 individuals, and one contained 43 individuals.

Laboratory Methods and Diagnostic Criteria

Plasma total cholesterol, triglycerides, LDL-B, LDL-C, VLDL-C, and apo E phenotype were determined. Plasma total cholesterol (Allain et al. 1974) and triglycerides (Sampson et al. 1975) were measured enzymatically, using an automated analyzer (ABA-100; Abbott, Pasadena). Apo B was measured using electroimmunoassay on the ultracentrifugal fraction with LDL density ($d > 1.006$ g/ml). The method of Reardon et al. (1981) was used but was varied by the addition of lipoprotein-deficient serum to the LDL standard (Rosseneu et al. 1981). LDL-C and VLDL-C content was measured enzymatically after separation of the lipoprotein fractions by the Lipid Research Clinic Protocol (U.S. Department of Health, Education, and Welfare 1974). The genotype at the locus coding for apo E was assigned based on the apo E phenotype determined by isoelectric focusing on a standard amount of delipidized VLDL (Bouthillier et al. 1983). Seven persons did not have apo E phenotypes measured, resulting in a sample size of 387 for all analyses involving the apo E genotype.

Statistical Analyses

Multiple linear regression was used to adjust LDL-B levels for variation associated with age and sex (B;as), age, sex, and apo E genotype (B;ase), or age, sex, apo E genotype, and LDL-C level (B;asec). The six apo E genotypes were coded as five 0, 1 categorical variables. Analysis of covariance was used to test for heterogeneity of the linear effects of age, sex, apo E genotypes, and LDL-C on LDL-B level among diagnostic categories (Sokal and Rohlf 1981). If such differences were found, separate adjustments of LDL-B levels for variation in age, sex, apo E genotype, and LDL-C level in each diagnostic category were carried out.

Investigation of the role of genetic and environmental influences in determining interindividual variability in LDL-B levels was carried out by fitting a general model and contrasting it with reduced models using complex segregation analysis (Elston 1986). The general model describes the distribution of LDL-B levels in terms of the independent contributions of variability of a single factor with a major effect, polygenes, and nongenetic effects specific to the individual. This distribution is formed by the commingling of three conditional distributions, or modes, associated with the occurrence of a major effect with two factors, $B\ell$ which lowers and $B\not{\ell}$ which raises the LDL-B level of the individual. Each individual has two doses of this factor, forming the combinations $B\ell B\ell$, $B\ell B\not{\ell}$, or $B\not{\ell}B\not{\ell}$, which are called ousiotypes (Cannings et al. 1978). Large differences in the mean LDL-B level among ousiotypes may result in the existence of three modes in the LDL-B distribution. These three conditional distributions for ousiotypes $B\ell B\ell$, $B\ell B\not{\ell}$, or $B\not{\ell}B\not{\ell}$ have means μ_1, μ_2, and μ_3, respectively. The ousiotype effects may be attributable to either a genetic or an environmental factor. Deviations from the ousiotype means are assumed to be attributable to the summation of effects of many unlinked polygenes, each with small additive effects, and individual-specific, nongenetic environmental effects. These deviations are assumed to be normally distributed. Environmental deviations are assumed to be independently distributed among relatives and to include errors of measurement. The major factor effects, the polygenic effects, and the individual-specific environmental effects are assumed to be uncorrelated and additive. The probabilities that a parent with ousiotype $B\ell B\ell$, $B\ell B\not{\ell}$, or $B\not{\ell}B\not{\ell}$ will transmit a $B\ell$ factor are τ_1, τ_2, or τ_3, respectively (Elston and Stewart 1971). The parameter p is defined as the relative frequency of factor $B\ell$, which is associated with the lowering of the LDL-B level phenotype, in the population being sampled. The relative frequency of the $B\not{\ell}$ factor is taken to be $q = 1 - p$. The relative frequencies of ousiotypes for individuals whose parents are not sampled are assumed to be those given by the Hardy-Weinberg distribution. They are p^2, $2pq$, and q^2 for the $B\ell B\ell$, $B\ell B\not{\ell}$, or $B\not{\ell}B\not{\ell}$, ousiotypes, respectively. The expected ousiotype frequencies for individuals whose parents are included in the sample are a function of the parents' ousiotypes and the transmission probabilities. The fraction of the variance among individuals with the same ousiotype (σ^2) which is attributable to polygenic loci is represented by the parameter h^2.

Submodels of the general model are defined by restricting the parameters found in the general model to specific values. The *major environmental model* restricts $\tau_1 = \tau_2 = \tau_3 = p$. Under this model there are three modes, but the a priori probability of an offspring having a certain ousiotype is independent of the parental ousiotypes. The *mixed codominant model* restricts the

transmission probabilities to the expected Mendelian segregation probabilities of a single locus with two alleles. Under this model the values of τ_1, τ_2, and τ_3 are 1.0, .5, and .0, respectively. The *major gene model* also restricts τ_1, τ_2, and τ_3 to 1.0, .5, and .0, and, in addition, it restricts h^2 to .0.

When the mixed codominant model fits as well as the general model but the major environmental model gives significantly worse fit, reduced genetic models are examined for evidence of dominance or recessiveness at the locus determining the major effect. By setting the phenotypic mean of the heterozygotes equal to that of the $B_{\not l}B_{\not l}$ homozygotes ($\mu_2 = \mu_3$), a *mixed dominant model,* where the heterozygote has the same genotypic value as a person homozygous for the $B_{\not l}B_{\not l}$ allele, is constructed. In the *mixed recessive model* the phenotypic mean of the heterozygotes is set equal to the mean of the $B_l B_l$ homozygotes ($\mu_2 = \mu_1$). The single-mode polygenic and sporadic models were fit for all hierarchies. The *polygenic model* is a single-mode model with deviations due to polygenic and individual-specific environmental effects. The *sporadic model* restricts $\mu_1 = \mu_2 = \mu_3$ and $h^2 = 0$. Under this model the total phenotypic variability is attributable entirely to environmental factors and measurement error specific to the individual.

Quasi-Newtonian, variable-metric approximation methods to determine the maximum likelihood estimates and standard errors of the parameter estimates were incorporated using GEMINI (Lalouel 1979). The model likelihoods on the pedigree data were approximated by the *p*edigree *a*nalysis *p*rogram (PAP) developed by Hasstedt (Hasstedt and Cartwright 1979; Hasstedt 1982). Hypothesis testing relied on use of the likelihood ratio test (LRT) to contrast pairs of models. Under the null hypothesis, and given a large sample size, the distribution of $-2(\log_e \kappa)$ is approximated by a χ^2 distribution, where κ equals the likelihood of the parameters in the reduced model given the data divided by the likelihood of the parameters in the full model given the data. The χ^2 distribution has df equal to the difference in the number of parameters between the reduced and full models. The order of comparison of models was hierarchical, beginning with the most complex model, the general model. The preferred model was taken to be the one with the smallest number of parameters that does not fit the data significantly worse than the next best-fitting model.

For a trait that is skewed, contrasting the mixed model with reduced models may lead to the false inference that a major locus exists (Maclean et al. 1975). However, normalizing transformations of a biologically skewed dependent variable can lead to a large reduction in the power to detect a major gene when one exists (Maclean et al. 1976). The consideration of the major environmental model as an alternative explanation for the commingling of multiple distributions reduces the possibility that skewness alone will lead to a false inference regarding the presence of a single locus with a large effect (Demenais et al. 1986). Should skewness be responsible for improved fit of the major gene model, the major environmental model should fit as well, or better, as it does not require the genetic transmission of the extreme values. Because of these considerations, no normalizing transformation was performed. Hasstedt et al. (1987) also chose not to perform a normalizing transformation in their segregation analysis of apo B levels. Hence, a more direct comparison of their results with those presented here is possible. No correction for ascertainment was made for two reasons. First, the population of inference is the sample ascertained through probands who were patients in the lipid clinic. Examination of the genetic control of LDL-B levels in a sample of families identified through lipid clinic patients is a primary motivation of this study. Second, implementation of an ascertainment correction strategy that would allow inferences about the population at large is technically limited by the variety of pedigree structures and by the fact that the pedigrees used in this study were ascertained through patients with a variety of disorders that may or may not be related to LDL-B level.

A two-factor analysis of variance (ANOVA) was used to test for heterogeneity, between diagnostic categories, of the effects of apo E genotypes and for allelic variation at the single locus with a large effect on LDL-B levels defined by the segregation analysis on LDL-B levels. Evaluation of the unmeasured single-locus effect required the assignment of a most probable LDL-B genotype. The assignment of an unmeasured genotype to each individual is accomplished by comparing the three genotypic probabilities. To calculate these three probabilities for each individual, complex segregation analysis is used to compute the likelihood that the individual has the genotype $B_l B_l$, the likelihood that the individual has the genotype $B_l B_{\not l}$, and the likelihood that the individual has the genotype $B_{\not l}B_{\not l}$. These three likelihoods are conditional on the genetic model and on the parameter estimates for the entire sample, the phenotype of the individual, and the distribution of phenotypes in the individual's family. The probability that the individual has a given genotype equals the likeli-

hood of that genotype divided by the sum of the three likelihoods. For these analyses all individuals had greater than a .50 probability of having one of the three genotypes and were assigned that genotype. Work by Odenheimer (1985) indicates that when data on at least five-member pedigrees (two parents, three offspring) are available, and when the effect of the major locus exceeds one within mode SD, both sensitivity and specificity of genotype assignment exceed .90 when one uses a predicted probability of .50 to define the presence of the genotype.

Results

Preliminary Analyses

Multiple regression analyses were performed to examine the effect of variation in age, sex, height, weight, apo E genotype, and LDL-C level on LDL-B level. Quadratic (squared) independent variables were considered for those quantitative independent variables whose linear term was significant. Age and sex differences explained a significant fraction (19.2%) of the variation in LDL-B level ($P < .0001$). Given age and sex, the addition of height, weight, and age squared did not explain a significant increase in the variation explained by the regression equation ($P > .05$). However, inclusion of apo E phenotypes into the regression model with linear age and sex did improve prediction of LDL-B level significantly ($P < .001$). Together, variation in age, sex, and apo E phenotype explained 37.3% of the variation in LDL-B level. The ε2 allele-containing genotypes are associated with a decrease in LDL-B level, and the ε4 allele-containing genotypes are associated with an increase in LDL-B level. These results are consistent with those reported by Sing and Davignon (1985) and Utermann (1985). The addition of LDL-C to the regression model also contributed significantly to the prediction of LDL-B level ($P < .001$). For the sample considered here, together age, sex, apo E genotype, and LDL-C explain 83.8% of the variation in LDL-B level.

The analysis of covariance was used to determine whether the regressions were homogeneous among diagnostic categories. Four diagnostic categories were considered: (1) normal, including subjects with high LDL-B levels in the presence of normal LDL-C and VLDL-C levels, (2) type IV HLP, (3) type III HLP, and (4) type IIa or IIb HLP. The type V HLP and unclassified mild hypertriglyceridemia classes were not included as diagnostic strata because they contained too few individuals. Homogeneity among strata of the regression of LDL-B level on age and sex, the regression of LDL-B level on age, sex, and apo E, or the regression of LDL-B level on age, sex, apo E, and LDL-C levels was not rejected ($P > .05$). These analyses confirm that separate adjustments of LDL-B, within each diagnostic category, for variation in age, sex, apo E genotype, and LDL-C level were not necessary in this sample. LDL-B level is not normally distributed in this sample. When the D'Agostino and Pearson (1973) criteria is used, the kurtosis of 2.906 is not significant at the .05 probability level. The skewness of .585 is significant ($P < .05$).

The results of the complex segregation analyses are presented in tables 1–4. Table 1 presents the analysis of LDL-B levels adjusted for variation in age and sex (B;as); table 2 gives the analysis of LDL-B levels adjusted for variation in age, sex, and apo E genotype (B;ase); and tables 3 and 4 give the analysis of LDL-B levels adjusted for variation in age, sex, apo E genotype, and LDL-C level (B;asec). The mixed codominant model does not give a significantly poorer fit than the general model, while the major environmental model is rejected for both the B;as (table 1) and B;ase (table 2) variables. Thus, the hypothesis that the transmission probabilities are Mendelian could not be rejected, while the hypothesis that the distribution of LDL-B is explained by an environmental factor causing a major effect is rejected.

We next consider submodels of the mixed codominant model. For B;as each of the submodels fits significantly worse than the mixed codominant model (table 1). Therefore, of the models considered, that containing a single locus with a large effect on LDL-B with polygenic and individual-specific environmental variation within genotype mean best fits these data. In the case of B;ase (table 2), all one- and two-mode submodels have significantly poorer fit than the mixed codominant model. However, the major-gene, three-mode model without a polygenic component did not fit significantly worse than the mixed codominant model at the .05 probability level. Thus, the hypothesis that the polygenic variance equals zero is not rejected after LDL-B levels have been adjusted for age, sex, and apo E genotype.

The effects of a major factor could not be detected using LDL-B adjusted for age, sex, apo E genotype, and LDL-C level (B;asec). The best-fitting three-mode models with polygenes gave estimates of the ousiotype means that were not ordered. That is, μ_2 did not fall between μ_1 and μ_3. This was interpreted as evidence that the three-mode model fitted these data poorly. Thus, only one- and two-mode models were considered for the B;asec variable. Separate comparisons of

Table 1

Complex Segregation Analysis: B;as in 83 Families, n = 394

	Model							
Parameter	General	Major Environmental	Mixed Codominant	Major Gene	Mixed Recombinant	Mixed Dominant	Polygene	Sporadic
μ_1	83.2	92.4	86.2	82.1	96.2	98.5	101.3	101.9
μ_2	127.4	141.9	129.5	126.4	(μ_1)	(μ_3)	. . .	. . .
μ_3	186.5	216.6	189.4	187.0	162.3	180.8	. . .	. . .
p	.828	.925	.826	.787	.734	.982	. . .	. . .
h^2	.512	.835	.463	. . .	.732	.646	.698	. . .
σ	21.05	25.93	23.14	21.27	28.54	29.92	33.22	34.23
τ_1	.963	(p)	(1.0)	(1.0)	(1.0)	(1.0)	. . .	. . .
τ_2	.423	(p)	(.5)	(.5)	(.5)	(.5)	. . .	. . .
τ_3	.000	(p)	(.0)	(.0)	(.0)	(.0)	. . .	. . .
−Likelihood ($\log_e$)	1890.0	1897.8	1891.7	1899.6	1896.8	1896.9	1908.0	1951.4
x^2 From general model		15.52*	3.32					
df From general model		3	3					
χ^2 From mixed codominant				15.71*	10.10*	10.42*	32.54*	119.40*
df From mixed codominant				1	1	1	3	4

NOTE.—Parentheses indicate that the value is fixed in the model.
* Significant at P = .05.

Table 2

Complex Segregation Analysis: B;ase in 83 Families, n = 387

	Model							
Parameter	General	Major Environmental	Mixed Codominant	Major Gene	Mixed Recombinant	Mixed Dominant	Polygene	Sporadic
μ_1	86.2	88.3	86.6	84.9	98.2	99.3	101.4	101.8
μ_2	124.1	123.6	121.8	120.0	(μ_1)	(μ_3)	. . .	. . .
μ_3	181.1	178.0	178.7	175.1	172.5	188.2	. . .	. . .
p	.836	.831	.798	.775	.794	.987	. . .	. . .
h^2	.398	.768	.282	. . .	.587	.526	.586	. . .
σ	19.63	20.48	20.25	19.28	25.98	26.96	30.03	30.10
τ_1	.958	(p)	(1.0)	(1.0)	(1.0)	(1.0)	. . .	. . .
τ_2	.433	(p)	(.5)	(.5)	(.5)	(.5)	. . .	. . .
τ_3	.000	(p)	(.0)	(.0)	(.0)	(.0)	. . .	. . .
−Likelihood ($\log_e$)	1823.8	1831.5	1827.0	1828.8	1830.8	1829.7	1844.1	1867.1
x^2 From general model		15.38*	6.38					
df From general model		3	3					
χ^2 From mixed codominant				3.65	7.64*	5.50*	34.26*	80.15*
df From mixed codominant				1	1	1	3	4

NOTE.—Parentheses indicate that the value is fixed in the model.
* Significant at P = .05.

Table 3

Complex Segregation Analysis: B;asec in 83 Families, $\mu_1 = \mu_2$, $n = 387$

Parameter	Model					
	General	Major Environmental	Mixed Recessive	Major Gene	Polygene	Sporadic
$\mu_1 = \mu_2$	95.8	96.9	97.0	94.7	102.2	102.3
μ_3	117.0	118.9	117.1	114.1	...	...
p	.487	.519	.489	.381	...	...
h^2	.653	.676	.363	...	.400	...
σ	12.17	12.36	12.77	12.09	15.38	15.37
τ_1	.896	(p)	(1.0)	(1.0)	...	...
τ_2	.124	(p)	(.5)	(.5)	...	...
τ_3	.602	(p)	(.0)	(.0)	...	...
−Likelihood ($\log_e$)	1591.4	1593.5	1595.1	1598.0	1596.3	1607.0
x^2 From general model		4.25	7.54	13.26*	9.75	31.23*
df From general model		3	3	4	6	7

NOTE.—Parentheses indicate that the value is fixed in the model.
* Significant at $P = .05$.

submodels for $\mu_1 = \mu_2$ (table 3) and $\mu_2 = \mu_3$ (table 4) were made. Neither the major environmental model nor the mixed genetic model fitted significantly worse than the two-mode general model, for both the $\mu_1 = \mu_2$ hierarchy ($\chi^2_3 = 4.25$ and 7.54, respectively) and the $\mu_2 = \mu_3$ hierarchy ($\chi^2_3 = 2.65$ and 4.92, respectively). The polygene model did not fit significantly worse than the general model for the case $\mu_1 = \mu_2$ or the case $\mu_2 = \mu_3$ ($\chi^2_6 = 9.75$ and 9.32, respectively). In both cases, the model without polygenes, the major-gene model ($\chi^2_4 = 13.26$ and 8.37, respectively), and sporadic model ($\chi^2_7 = 31.23$ and 30.80, respectively) gave a significantly worse fit than the general model. Because it has fewer parameters than the major environmental and mixed genetic models yet does not fit significantly worse, the polygene model is considered to be the best model for B;asec.

The best estimates of the effects of the genotypes at the single locus on LDL-B levels are given by the analysis of B;ase. The further reduction in variation around

Table 4

Complex Segregation Analysis: B;asec in 83 Families, $\mu_2 = \mu_3$, $n = 387$

Parameter	Model					
	General	Major Environmental	Mixed Dominant	Major Gene	Polygene	Sporadic
μ_1	96.4	97.0	94.4	93.8	102.2	102.3
$\mu_2 = \mu_3$	117.7	119.1	113.8	113.2	...	...
p	.850	.881	.768	.751	...	...
h^2	.663	.674	.238	...	.400	...
σ	12.39	12.41	12.21	12.01	15.38	15.37
τ_1	.844	(p)	(1.0)	(1.0)	...	...
τ_2	.999	(p)	(.5)	(.5)	...	...
τ_3	.009	(p)	(.0)	(.0)	...	...
−Likelihood ($\log_e$)	1591.6	1593.0	1594.1	1595.8	1596.3	1607.0
x^2 From general model		2.65	4.92	8.37*	9.32	30.80*
df From general model		3	3	4	6	7

NOTE.—Parentheses indicate that the value is fixed in the model.
* Significant at $P = .05$.

Table 5

Sources of the Variation in LDL-B Levels[a]

Effect	% Variation of B;ase	% Unadjusted LDL-B Variation
Age and sex	...	19.17
Apo E genotype, given age and sex	...	18.11
Single locus (V_{mg})	55.8	34.99[b]
Polygenic (V_{pg})	12.5	7.82[b]
Individual-specific environment (V_e)	31.7	19.91[b]

[a] The total phenotypic variance of B;ase is defined as $V_p = V_m + V_{pg} + V_e$, where V_m is the variance attributable to differences among ousiotypes, V_{pg} is the variance associated with polygenes, and V_e is the variance associated with individual-specific environmental and measurement error differences. In this model $V_m = \Sigma_j f_j (\mu_j - \mu..)^2$, where $\mu..$ is the mean LDL-B level of the sample and f_j and μ_j are the relative frequency and mean, respectively, of the jth ousiotype. $V_{pg} = h^2\sigma^2$. $V_e = \sigma^2 - V_{pg}$. The proportion of the total phenotypic variability due to the variance among ousiotype means = V_m/V_p. The proportions due to polygenic and environmental effects are V_{pg}/V_p and V_e/V_p, respectively. These variances are calculated by the same equations for the mixed codominant and major environmental models, where the single-locus effect and the major environmental effect replace the general model major effect. V_{mg} is the variance due to the single locus with a major effect, V_{me} is the variance due to the environmental major effect, and both are calculated identically to V_m.

[b] Given age, sex, and apo E genotype.

the mean of each genotype that is obtained by correction for age, sex, and apo E genotype versus correction for age and sex alone allows a more accurate estimate of the genotypic means. Therefore, the parameter estimates for the mixed codominant model for B;ase were used to calculate the percent of variation in LDL-B due to the unmeasured single locus, unmeasured polygenes, and individual-specific environments. These estimates are shown in table 5. One-third of the interindividual variability in apo B levels is explained by the segregation of a single locus with a major effect.

The magnitudes of the effects of apo E genotypes on LDL-B level adjusted for variation in age and sex for each diagnostic category are shown in table 6. The magnitude of the effects of LDL-B genotype on LDL-B level adjusted for age, sex, and apo E genotype can be found by comparing the estimates of the genotype means of the mixed codominant model in table 2. That these effects are homogeneous across diagnostic category is confirmed by the lack of statistical significance of the genotype by diagnostic-category interaction term in separate two-factor ANOVAs for apo E genotype in normals versus type IV HLP ($P = .85$) and in normals versus type II HLP when using only apo E genotypes that had more than one case in each diagnostic category ($P = .83$). Homogeneity is also found for the effects

Table 6

LDL-B Means by Apo E Genotype and Diagnostic category

	Apo E					
Diagnostic Category	ε2/2	ε3/2	ε4/2	ε3/3	ε4/3	ε4/4
Normal	54.98 (7)	78.94 (53)	91.05 (10)	99.91 (105)	111.49 (38)	96.09 (4)
Type IV HLP	83.50 (3)	96.56 (36)	113.79 (12)	113.18 (43)	121.63 (20)	104.67 (6)
Type II HLP	... (0)	... (0)	... (0)	166.62 (15)	180.17 (7)	184.23 (2)
Type III HLP	62.46[a] (18)					

Note.—Values in parentheses are cell sizes.

[a] Not included in the ANOVA but presented here for comparison.

of LDL-B genotypes across diagnostic categories (data not shown).

Discussion

Analysis of both LDL-B level adjusted for variation due to age and sex (B;as) and LDL-B level adjusted for variation due to age, sex, and apo E genotype (B;ase) detected evidence for a codominant single locus with a large effect on LDL-B level. The acceptance of this three-mode genetic model, while the three-mode environmental model is rejected, is strong evidence for the presence of segregation at a single locus with a large effect on LDL-B level. The effects of allelic variations at this locus appear to be homogeneous among diagnostic categories. The mean LDL-B;ase level for individuals with one B$\not\ell$ allele is approximately 35 mg/dl higher than for those without the B$\not\ell$ allele, and those with two copies of the B$\not\ell$ allele have a mean LDL-B level approximately 55 mg/dl higher than the heterozygotes. This major genetic factor explains 34.99% of the total phenotypic variation in unadjusted LDL-B levels in this sample. Polygenic effects explain 7.82%. The effects of variation in age, sex, and apo E genotype account for 37.28%, while the remaining 19.91% is attributable to individual-specific environmental effects and measurement errors.

Several lines of evidence argue that the locus detected by complex segregation analysis is not the gene coding for apo E. When the parameter estimates of the mixed codominant models for B;as and B;ase are contrasted, it is evident that the additional adjustment for apo E phenotype causes a large decrease in the estimate of the polygenic component of variation yet does not affect the detection of a large single-gene effect. The proportion of variation about the mean of each of the major-locus genotypes that is due to polygenes drops from .463 to .282 when the data are adjusted for differences among apo E genotype means. This difference of .181 is a drop of 39.1 percent. On the other hand, the estimates for the relative frequency of Bℓ allele, p, and the single-locus genotype means μ_1, μ_2, and μ_3 change less dramatically. The estimate of the variance associated with the single gene was 53.6% using B;as and 55.8% using the B;ase variable. Therefore, we conclude that the single-gene effect detected by these analyses is not associated with allelic variation at the apo E locus.

There are two possible explanations for the fact that, after correction for variation attributable to apo E genotypes, the proportion of within-genotype variability due to the polygenic effects was .282, not significantly different from zero at the .05 level of probability. First, in this sample there may be no other genetic factors which influence LDL-B level besides variation in sex, apo E genotype, and LDL-B genotype. The second alternative is that this sample may not provide enough power to reject the null hypothesis that $h^2 = 0$. Significant polygenic variation was detected in LDL-B adjusted for age, sex, apo E genotype, and LDL-C levels. Therefore, the existence of other genetic factors that are associated with variation in LDL-B levels cannot be excluded.

The results of the complex segregation analysis on LDL-B levels found here are in reasonable agreement with those of Hasstedt et al. (1987) for total apo B levels adjusted for variation in age and sex. The correlation of LDL-B level and total apo B level in the sample reported here is .78, allowing comparison of the apo B and LDL-B results. Hasstedt et al. detected a codominant locus for B;as with f(Bℓ) = .847 under the mixed codominant model, versus f(Bℓ) = .787 (for B;as) reported here. This is in reasonable agreement, since the pedigrees analyzed here were ascertained through a lipid clinic while only a portion of the Hasstedt et al. pedigrees were unrepresentative of the population at large. Ascertainment through a lipid clinic is expected to decrease p, since the proportion of affected patients carrying the lowering allele is expected to be lower than the proportion of normolipidemic individuals carrying the allele. Hasstedt et al. found that the separation between the low and middle means was 31.4 mg/dl and between the middle and high means was 66.2, as compared with 43.3 and 59.9 found for B;as here. These findings are also in reasonable agreement. The estimate of the polygenic component of within-genotype variability (h^2) of 15.3% found by Hasstedt et al. was not significantly different than zero, whereas the substantially larger estimate of 46.3% found for B;as here is significantly different from zero. However, adjustment for apo E genotype resulted in the estimate of the polygenic component of within-genotype variation decreasing to 28.2% in this study, which was not significantly different from zero. Though Hasstedt et al. did not have apo E genotypes available, it is likely that they had fewer individuals with non-ε3/3 genotypes because they had relatively fewer pedigress identified through a hyperlipidemic proband. This could account for the lower estimate of polygenic variability in B;as found by Hasstedt et al. The magnitude of the decline in h^2 and the resulting elimination of the significance of the h^2 term when LDL-B levels are adjusted for variation in apo E genotype, as well as for variation in age and sex, pro-

vide evidence for the large effect of variation in the apo E genotype on the polygenic component of LDL-B variability.

Beaty et al. (1986) did not detect statistically significant evidence for a single locus with a large effect on LDL-B levels adjusted for variation in age and sex. It is likely that their sample of one large Amish pedigree segregating for a lipid metabolic disorder (sitosterolemia) was not comparable with the samples examined here and by Hasstedt et al. Because of the Amish population's isolation from mating with other groups, the $B\not{b}$ allele may not exist in this population or simply not be segregating in the single large pedigree studied. Also, the existence of sitosterolemia may cause variation in LDL-B levels, resulting in decreased power to detect other genetic effects.

Evidence for a single-locus effect on LDL-B levels was removed by adjustment for LDL-C. This supports the hypothesis that the genetic polymorphism detected for LDL-B levels affects both LDL-B and LDL-C levels, as indeed it must given that apo B and cholesterol are both major constituents of LDL. Teng et al. (1983) and Sniderman et al. (1985) have shown that while the amount of cholesterol per LDL particle can vary, the quantity of the apo B per particle is fixed and that therefore LDL B is a more accurate measurement of the LDL particle number than is LDL cholesterol. Even so, because cholesterol remains a quantitatively important constituent of all LDL particles, there must inevitably be a close correlation between LDL-B and LDL cholesterol levels. These results are consistent with those of Hasstedt et al. (1987), who found a significant difference in LDL-C levels among individuals with differing apo B genotypes.

Conclusions

The results of complex segregation analyses provide evidence for the existence of a single locus with a large effect on interindividual variation in the level of LDL-B in this sample. Further, the continued detection of the effects of this locus after adjustment for genotypic variability at the locus coding for the apo E protein provides convincing evidence that the locus detected is not the apo E locus. This is the first report of a single locus with a large effect on LDL-B levels that is shown to be separate from the effect of the apo E polymorphism on LDL-B levels. The apo E polymorphism may be responsible for a major fraction of the polygenic component of interindividual variability in LDL-B levels.

Acknowledgments

This research was supported by grants from the Medical Research Council of Canada (MA-5427), the Ministry of Health and Welfare of Canada (NHRDP 6605-2087-52), and the National Institutes of Health (HL 24489 and HL 39107).

References

Allain, C. C., L. S. Poon, F. C. S. Chan, W. Richmond, and C. P. Fu. 1974. Enzymatic determination of total serum cholesterol. Clin. Chem. **20**:470–475.

American Heart Association. 1984. Recommendations for the treatment of hyperlipidemia in adults: a joint statement of the Nutrition Committee and the Council on Arteriosclerosis of the American Heart Association. Arteriosclerosis **4**:445A–468A.

Amos, C. I., R. C. Elston, S. R. Srinivasan, A. F. Wilson, J. L. Cresanta, L. J. Ward, and G. S. Berenson. 1987. Linkage and segregation analyses of apolipoproteins A1 and B, and lipoprotein cholesterol levels in a large pedigree with excess coronary heart disease. Genet. Epidemiol. **4**:115–128.

Avogaro, P., B. G. Bittolo, G. Cazzolato, and G. B. Quinci. 1979. Are apolipoproteins better discriminators than lipids for atherosclerosis? Lancet **1**:901–907.

Avogaro, P., B. G. Bittolo, G. Cazzolato, G. B. Quinci, and F. Belussi. 1978. Plasma levels of apolipoprotein AI and apolipoprotein B in human atherosclerosis. Artery **4**: 385–394.

Beaty, T. H., P. O. Kwiterovich, Jr., M. J. Khoury, S. White, P. S. Bachorik, H. H. Smith, and A. Sniderman. 1986. Genetic analysis of plasma sitosterol, apoprotein B and lipoproteins in a large Amish pedigree with sitosterolemia. Am J. Hum. Genet. **38**:492–504.

Beaumont, J. L., L. A. Carlson, G. R. Cooper, Z. Fejfar, D. S. Fredrickson, and T. Strasser. 1970. Classification of hyperlipidaemias and hyperlipoproteinaemias. Bull. WHO **43**: 891–908.

Bouthillier, D., C. F. Sing, and J. Davignon. 1983. Apolipoprotein E phenotyping with a single gel method: application to the study of informative matings. J. Lipid Res. **24**:1060–1069.

Cannings, S. C., E. A. Thompson, and M. H. Skolnick. 1978. Probability functions on complex pedigrees. Adv. Appl. Probability **10**:26–61.

D'Agostino, R., and E. S. Pearson. 1973. Tests for departure from normality: empirical results for the distributions of $\sqrt{b_1}$ and b_2. Biometrika **60**:613–622.

Demenais, F., M. Lathrop, and J. M. Lalouel. 1986. Robustness and power of the unified model in the analysis of quantitative measurements. Am J. Hum. Genet. **38**:228–234.

Elston, R. C. 1986. Modern methods of segregation analysis. Pp. 213–224 *in* S. H. Moolgavkar, and R. L. Prentice, eds. Modern statistical methods in chronic disease epidemiology. John Wiley & Sons, New York.

Elston, R. C., and J. Stewart. 1971. A general model for the genetic analysis of pedigree data. Hum. Hered. **21**:523–542.

Hasstedt, S. 1982. A mixed model likelihood approximation on large pedigrees. Comp. Biomed. Res. **15**:295–307.

Hasstedt, S., and P. Cartwright. 1979. PAP: pedigree analysis package. Tech. rep. 13. Department of Medical Biophysics and Computing. University of Utah, Salt Lake City.

Hasstedt, S. J., L. Wu, and R. R. Williams. 1987. Major locus inheritance of apolipoprotein B in Utah pedigrees. Atheroscler. Genet. Epidemiol. **4**:67–76.

Heiss G., I. Tamir, C. E. Davis, H. A. Tyroler, B. M. Rifkind, G. Schonfeld, D. Jacobs, and I. D. Frantz. 1980. Lipoprotein-cholesterol distributions in selected North American populations: the Lipid Research Clinics Program Prevalence Study. Circulation **61**:302–315.

Lalouel, J. M. 1979. GEMINI—a computer program for optimization of general nonlinear functions. Tech. rep. 14. Department of Medical Biophysics and Computing, University of Utah, Salt Lake City.

Maclean, C. J., N. E. Morton, R. C. Elston, and S. Yee. 1976. Skewness in commingled distributions. Biometrics **32**: 695–699.

Maclean, C. J., N. E. Morton, and R. Lew. 1975. Analysis of family resemblance. IV. Operational characteristics of segregation analysis. Am. J. Hum. Genet. **27**:365–384.

Odenheimer, D. J. 1985. An evaluation of the validity of complex segregation analysis in identifying an individuals genotype at a major locus. Ph.D. diss., University of Michigan, Ann Arbor.

Reardon, M. F., M. E. Poapst, K. D. Uffelman, and G. Steiner. 1981. Improved method for quantification of B apoprotein in plasma lipoproteins by electroimmunoassay. Clin. Chem. **27**:892–895.

Riesen, W. F., R. Mordasini, C. Salzmann, A. Theler, and H. P. Gurtner. 1980. Apoproteins and lipids as discriminators of severity of coronary heart disease. Atherosclerosis **37**:157–162.

Rosseneu, M., N. Vinaimont, R. Vercaemst, W. Dekeersgieter, and F. Belpaire. 1981. Standardization of immunoassays for the quantification of plasma Apo B protein. Anal. Biochem. **116**:204–210.

Sampson, E. J., L. M. Demers, and A. F. Krieg. 1975. Faster enzymatic procedure for serum triglycerides. Clin. Chem. **21**:1983–1985.

Schmidt, S. B., A. G. Wasserman, R. A. Muesing, S. E. Schlesselman, J. C. Larosa, and A. M. Ross. 1985. Lipoprotein and apolipoprotein levels in angiographically defined coronary atherosclerosis. Am. J. Cardiol. **55**:1459–1462.

Sing, C. F., and J. Davignon. 1985. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation. Am. J. Hum. Genet. **37**:268–285.

Sniderman, A., S. Shapiro, D. Marpole, B. Skinner, B. Teng, and P. O. Kwiterovich, Jr. 1980. Association of coronary atherosclerosis with hyperapobetalipoproteinemia [increased protein but normal cholesterol levels in human plasma low density (β) lipoproteins]. Proc. Natl. Acad. Sci. USA **77**:604–608.

Sniderman, A., B. Teng, J. Genest, K. Cianflone, S. Wacholder, and P. Kwiterovich, Jr. 1985. Familial aggregation and early expression of hyperapobetalipoproteinemia. Am. J. Cardiol. **55**:291–295.

Sniderman, A. D., C. Wolfson, B. Teng, F. A. Franklin, P. S. Bachorik, and P. O. Kwiterovich, Jr. 1982. Association of hyperapobetalipoproteinemia with endogenous hypertriglyceridemia and atherosclerosis. Ann. Int. Med. **97**: 833–839.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. 2d ed. W. H. Freeman, New York.

Teng, B., G. R. Thompson, A. D. Sniderman, T. M. Forte, R. M. Krauss, and P. O. Kwiterovich, Jr. 1983. Composition and distribution of low density lipoprotein fractions in hyperapobetalipoproteinemia, normolipidemia, and familial hypercholesterolemia. Proc. Natl. Acad. Sci. USA **80**:6662–6666.

U. S. Department of Health, Education and Welfare. 1974. Manual of laboratory operations. Lipid Research Clinics Program publ. (NIH) 75–628. U.S. Department of Health, Education and Welfare, Washington, DC.

Utermann, G. 1985. Genetic polymorphism of apolipoprotein E—impact on plasma lipoprotein metabolism. Pp. 1–28. *in* G. Grepaldi, A. Tiengo, and G. Baggio, eds. Diabetes, obesity and hyperlipidemias—III. Elsevier Science Publishers B.V. (Biomedical Division).

Whayne, T. F., P. Alaupovic, M. D. Curry, E. T. Lee, P. S. Anderson, and E. Schechter. 1981. Plasma apolipoprotein B and VLDL-, LDL-, and HDL-cholesterol as risk factors in the development of coronary artery disease in male patients examined by angiography. Atherosclerosis **39**: 411–424.