Comparative analysis of the apo(a) gene, apo(a) glycoprotein, and plasma concentrations of Lp(a) in three ethnic groups. Evidence for no common "null" allele at the apo(a) locus.

A Gaw, E Boerwinkle, J C Cohen, H H Hobbs

J Clin Invest. 1994;93(6):2526-2534. https://doi.org/10.1172/JCI117263.

Research Article

Distributions of plasma lipoprotein(a) (Lp[a]) concentrations exhibit marked interracial differences. Apolipoprotein(a) (apo[a]), the unique constituent of Lp(a), is highly polymorphic in length due to allelic variations in the number of kringle 4(K-4)-encoding sequences. Plasma Lp(a) concentrations are inversely related to the number of K-4 repeats in the apo(a) alleles. To determine the contribution of this length variation to the interracial variation in plasma Lp(a) levels, we compared apo(a) allele size, glycoprotein size, and plasma Lp(a) concentrations in Caucasians, Chinese, and African Americans. Caucasians and African Americans had very different distributions of plasma Lp(a) concentrations yet there was no significant difference in the overall frequency distributions of their apo(a) alleles. Over the entire size spectrum of apo(a) alleles, the plasma Lp(a) levels were higher in African Americans than in Caucasians. Conversely, Caucasians and Chinese had similar plasma Lp(a) concentrations but significantly different apo(a) allele size distributions. Therefore, interracial differences in the plasma concentrations of Lp(a) are not due to differences in the frequency distributions of apo(a) alleles. We also examined the relationship between apo(a) allele size and the presence of detectable plasma apo(a) protein in plasma. Apo(a) alleles associated with no detectable plasma protein were not of uniformly large size, as had been expected, but were distributed over the entire size spectrum. From this analysis, we conclude that there is no […]

Comparative Analysis of the *Apo(a)* Gene, Apo(a) Glycoprotein, and Plasma Concentrations of Lp(a) in Three Ethnic Groups

Evidence For No Common "Null" Allele at the Apo(a) Locus

Allan Gaw,* Eric Boerwinkle,‡ Jonathan C. Cohen,§ and Helen H. Hobbs*

**Departments of Molecular Genetics and Internal Medicine and §Center for Human Nutrition, University of Texas Southwestern Medical Center, Dallas, Texas 75235; and ‡Center for Demographic and Population Genetics, University of Texas Health Science Center, Houston, Texas 77225*

Abstract

Distributions of plasma lipoprotein(a) (Lp[a]) concentrations exhibit marked interracial differences. Apolipoprotein(a) (apo[a]), the unique constituent of Lp(a), is highly polymorphic in length due to allelic variations in the number of kringle 4(K-4)-encoding sequences. Plasma Lp(a) concentrations are inversely related to the number of K-4 repeats in the apo(a) alleles. To determine the contribution of this length variation to the interracial variation in plasma Lp(a) levels, we compared apo(a) allele size, glycoprotein size, and plasma Lp(a) concentrations in Caucasians, Chinese, and African Americans. Caucasians and African Americans had very different distributions of plasma Lp(a) concentrations yet there was no significant difference in the overall frequency distributions of their apo(a) alleles. Over the entire size spectrum of apo(a) alleles, the plasma Lp(a) levels were higher in African Americans than in Caucasians. Conversely, Caucasians and Chinese had similar plasma Lp(a) concentrations but significantly different apo(a) allele size distributions. Therefore, interracial differences in the plasma concentrations of Lp(a) are not due to differences in the frequency distributions of apo(a) alleles.

We also examined the relationship between apo(a) allele size and the presence of detectable plasma apo(a) protein in plasma. Apo(a) alleles associated with no detectable plasma protein were not of uniformly large size, as had been expected, but were distributed over the entire size spectrum. From this analysis, we conclude that there is no common "null" allele at the apo(a) locus. (*J. Clin. Invest.* 1994. 93:2526–2534.) Key words: apolipoprotein(a) • lipoprotein(a) • polymorphism • Caucasian • Chinese • African American

Introduction

Lipoprotein(a) (Lp[a]),[1] a cholesterol ester-rich lipoprotein, is a derivative of low density lipoprotein (LDL), with distinct structural, epidemiological, and genetic properties (1, 2). Lp(a) contains a particle of LDL to which is attached a large, hydrophilic glycoprotein, apolipoprotein(a) (apo[a]). Plasma levels of Lp(a) vary over a 1000-fold range among individuals in a population but remain remarkably constant within an individual over time (3). In both Caucasian and Oriental populations, elevated plasma levels of Lp(a) confer a 1.5–3-fold increase in the risk of coronary atherosclerosis (4–7). Because of the association between high plasma Lp(a) concentrations and coronary artery disease, the factors controlling plasma Lp(a) concentrations have been extensively studied over the last five years.

Soon after the discovery of Lp(a), family studies suggested that its plasma levels were highly heritable (8). The apo(a) gene was implicated as a major determinant of plasma Lp(a) levels when it was noted that the apo(a) glycoprotein varied in size over a wide range and that its size tended to be inversely related to plasma concentrations of Lp(a) (1, 9). The molecular mechanism responsible for this size variation is length polymorphism in the *apo(a)* gene (10). The *apo(a)* gene contains several tandem repeats of a 5.5-kb sequence encoding a cysteine-rich protein motif resembling the kringle 4 (K-4) of plasminogen (11). The number of K-4 domain repeats in different apo(a) alleles varies from 12 to 51 (12). Approximately 94% of Caucasians have two apo(a) alleles that contain different numbers of K-4 repeats (10). This highly informative length polymorphism has been used in family studies to examine the relationship between apo(a) genotype and plasma concentrations of Lp(a). It has been estimated that in Caucasians, ~ 90% of the inter-individual variation in plasma Lp(a) levels is explained by sequence differences at, or closely linked to, the apo(a) locus (13). To date, this finding has not been extended to other racial groups, which is important since the distributions of plasma Lp(a) levels exhibit marked interracial variability.

In Caucasians and Chinese, the distribution of plasma Lp(a) is highly skewed towards lower levels (2, 4, 14–16). However, in Africans and African Americans, the distribution curve of plasma Lp(a) levels has a more Gaussian shape (1, 14–20) and on average, African Americans have 2–3-fold higher plasma concentrations of Lp(a) than either Caucasians or Chinese. The factors responsible for the racial differences in distribution of Lp(a) levels have not been identified, but previous studies have suggested that differences in apo(a) isoform size distributions are not solely responsible (14–16, 18). The use of apo(a) isoforms in such studies is problematic, however, because standard immunoblotting methods cannot detect protein products from all apo(a) alleles. In fact, in only one out of eleven ethnic groups in which apo(a) allele frequencies have been calculated were the alleles in Hardy-Weinberg equilibrium (16, 18).

Address correspondence to Helen H. Hobbs, Department of Molecular Genetics, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-9046.

Received for publication 1 November 1993 and in revised form 17 February 1994.

1. Lp(a), lipoprotein(a); apo(a), apolipoprotein(a); LDL, low-density lipoprotein; HDL, high-density lipoprotein; TG, triglyceride; K-4, kringle-4; kb; kilobases; EDTA, ethylenediamine tetraacetic acid; ELISA, enzyme-linked immunosorbent sandwich assay; SDS, sodium dodecyl sulfate.

J. Clin. Invest.

0021-9738/94/06/2526/09 $2.00
Volume 93, June 1994, 2526–2534

In this study, we used high resolution apo(a) genotyping and phenotyping systems (12) to evaluate the relationship between apo(a) gene structure, apo(a) glycoprotein size, and plasma concentrations of Lp(a) in three racial groups: Caucasians, Chinese, and African Americans. This is the first study in which apo(a) gene structure, apo(a) glycoprotein size, and plasma Lp(a) levels have been simultaneously analyzed in any population.

Methods

Plasma and lipoprotein analysis. Fasting blood samples were collected into Na_2 EDTA tubes from a total of 381 unrelated individuals residing within 50 miles of Dallas, TX. Plasma was isolated by low speed centrifugation and stored at −70 °C. Plasma Lp(a) concentrations were measured within 2 mo of collection at GeneScreen (Dallas, TX) using an enzyme-linked immunosorbent sandwich assay (ELISA) (21). The capture antibody used in the assay was a polyclonal rabbit anti-human Lp(a) antibody and the detection antibody was a mouse monoclonal anti-human apo(a) antibody, IgG-$1A^2$, provided by Dr. Gerd Utermann (University of Innsbruck, Innsbruck, Austria). Total cholesterol and triglyceride levels and plasma lipoprotein concentrations were assayed on fresh plasma in the laboratory of Dr. Scott Grundy (UT Southwestern Medical Center, Dallas, TX) according to the protocols of the Lipid Research Clinics (22).

Apo (a) isoform determination by immunoblotting. Apo (a) isoform analysis was performed using a modification of the immunoblotting method described by Kamboh et al. (23). A 12.5-μl aliquot of plasma was mixed with 30 μl reducing buffer (1:2:10 ratio of β-mercaptoethanol, 0.5% [wt/vol] bromophenol blue in 5% [vol/vol] glycerol, 5% [wt/vol] SDS) and boiled for 5 min. A 10-μl aliquot was loaded onto a 2% agarose gel (15 × 25 cm) in 90 mM Tris-HCl, 90 mM boric acid, 2 mM Na_2EDTA and 0.1% (wt/vol) SDS. Electrophoresis was performed in a horizontal gel apparatus (BRL Life Technologies, Gaithersburg, MD) in 45 mM Tris-HCl, 45 mM boric acid, 2 mM Na_2EDTA and 0.1% (wt/vol) SDS for 17 h at 120 V at 4°C. After electrophoresis, proteins were transferred to a supported nitrocellulose membrane (Hybond™-C extra, Amersham, Arlington Heights, IL) by electroblotting using a Hoefer Transphor Cell (Hoefer Scientific Instruments, San Francisco, CA) at 90 V for 3 h in 10 mM Tris-HCl, 40 mM glycine and 5% (vol/vol) methanol at 4 °C. The membrane was incubated in TBS (0.5 M NaCl and 35 mM Tris-HCl, pH 7.4), 5% (wt/vol) powdered skimmed milk and 0.2% (vol/vol) Tween 20 for 1 h before a 30-min incubation in TBS plus 1% (vol/vol) NP-40 and 400 μg/ml mouse monoclonal anti-human apo(a)-specific antibody IgG-$1A^2$ (21), which had been conjugated to horseradish peroxidase (24). IgG-$1A^2$ is directed against the common K-4 repeat, i.e., the type A and type B K-4 repeat, (reference 11 and Dr. Gerd Utermann, personal communication). The membrane was washed four times in 100 ml washing buffer (TBS plus 1% [vol/vol] NP-40, 0.5% [wt/vol] deoxycholic acid and 0.1% [wt/vol] filtered SDS) and then developed by using ECL™ Western blot detection reagent (Amersham) and subjected to autoradiography using X-OMAT™ AR film (Eastman Kodak Co., Rochester, NY) for 1–10 min.

To determine the sensitivity of the assay, serial dilutions of plasmas with apo(a) alleles of various size were examined by immunoblotting. All isoforms with plasma Lp(a) concentration of ≥ 0.1 mg/dl were detected (data not shown). The sensitivity of the immunoblotting detection system was the same irrespective of apo(a) isoform size. In addition, on a subset of 15 plasma samples, we also performed immunoblotting using a monoclonal anti-peptide antibody (10C5) directed against an epitope in the protease domain of the apo(a) glycoprotein. Both antibodies gave signals of equal relative intensity and no isoforms that were undetectable by $1A^2$ were revealed using 10C5.

Apo(a) genotype determination by pulsed-field gel electrophoresis and genomic blotting Apo(a) alleles were examined using a modification of the pulse-field gel electrophoresis procedures described by Lackner et al. (10). Leukocytes were isolated from whole blood using LeucoPrep tubes (Becton, Dickinson & Co., Lincoln Park, NJ), embedded in agarose plugs, and stored at 4°C in 0.5 M Na_2 EDTA, pH 8.0. A 3-mm segment of the agarose–cellular plug was washed in 1 × TE (1 mM Na_2 EDTA, 10 mM Tris-HCl, pH 8.0) and incubated twice for 2 h each with 100 U KpnI or HpaI (New England Biolabs, Beverly, MA) in 70 μl of the buffer suggested by the manufacturer (10, 12). The plugs were subjected to transverse alternating gel electrophoresis using a GeneLine II™ apparatus (Beckman Instruments, Inc., Fullerton, CA). The pulse times were 4 s for 30 min followed by 6-s pulses (150 V/350 mAmps) for 18 h. Low range PFG markers (New England Biolabs) were used as size standards. The size-fractionated DNA was blotted to a nylon membrane (ICN Biomedicals, Irvine, CA) and hybridized with an [α-^{32}P]dCTP radiolabeled human apo(a) K-4–specific single-stranded fragment (10). The filters were washed for 1 h and exposed to Kodak X-OMAT™ AR film with an intensifying screen (Lightening Plus Dupont Co., Wilmington, DE) for 16 h at −70°C. The apo(a) alleles were designated by the estimated number of K-4–encoding sequences per allele (*apo(a) K-12* to *apo(a) K-51*) (12).

Statistical methods. The overall shape of the distribution of plasma Lp(a) concentrations was compared between populations using the two-sample Kolmogorov-Smirnov test (25). The distributions of plasma Lp(a), were clearly not normally distributed, therefore, nonparametric rank-sum tests and Kruskal-Wallis tests (25) were used to compare median Lp(a) levels between racial groups.

Because of the large number of apo(a) phenotypes relative to the numbers of individuals in this sample, traditional chi-square goodness-of-fit test and chi-square tests of independence were inappropriate. A goodness-of-fit test to Hardy-Weinberg expectations was carried out using a computerized permutation algorithm described by Guo and Thompson (26). Comparisons of apo(a) isoform and phenotype frequencies among races were carried out using a Monte Carlo simulation method to derive P values for Fisher's exact test (27).

The relationship between the apo(a) polymorphism and plasma Lp(a) concentrations was analyzed in several ways. The contribution of the apo(a) genotypes to the inter-individual variance in plasma Lp(a) concentrations was estimated using the adjusted R^2 from an analysis of variance (28). The Spearman's rank correlation (25) between the sum of the number of K-4 repeats represented in an individual's genotype and their plasma Lp(a) level was calculated. These rank correlations were compared between populations using Fisher's z-transformation.

Results

Venous blood was obtained from 174 unrelated American Caucasians (86 men, 88 women), 101 unrelated Chinese (46 men, 55 women), and 106 unrelated African-Americans (20 men, 86 women). All subjects were selected at random from healthy individuals living in the Dallas–Fort Worth area. The only selection criteria were that all Caucasians and African Americans had to be born in the United States and all Chinese had to be born in Mainland China or Taiwan. Some data from 103 of the 174 Caucasians were included in a prior publication (10). Four Caucasians, 17 Chinese and 12 African Americans were postmenopausal at the time of sampling. A description of the characteristics of the three samples is given in Table I. There was no significant difference in the ages, or plasma levels of total cholesterol or LDL-cholesterol(C) of the three groups. The African Americans had slightly lower plasma concentrations of triglyceride and higher concentrations of high density lipoprotein HDL(C) than the other two racial groups.

The distribution of plasma Lp(a) concentrations in each of the three racial groups is shown in Fig. 1. The characteristics of the distributions of plasma Lp(a) levels were compared in the three populations (Table II). In all populations, the plasma Lp(a) concentrations varied over a wide range. As reported previously, (14–16, 18) the plasma levels of Lp(a) were more

Table I. Description of the Samples from Each of the Three Populations

Race	n	Age	Plasma levels (mg/dl) Chol.*	Triglyceride	LDL-C	HDL-C
		yr				
Caucasian	174	38±9	195±62	130±93	119±34	50±14
Chinese	101	44±16	190±34	141±108	117±30	51±13
African American	106	39±10	195±39	100±56	126±37	55±13

* All values are given as means±SD.

highly skewed towards the lower range in the Chinese (skew = 3.77) and Caucasians (skew = 1.33) than in the African Americans (skew = 0.84). There was only a marginal statistical difference between the overall distribution of plasma Lp(a) levels (Kolmogorov-Smirnov, P = 0.06) between the Chinese and Caucasians; the distributions in both populations were significantly different from the African Americans. (P < 0.001).

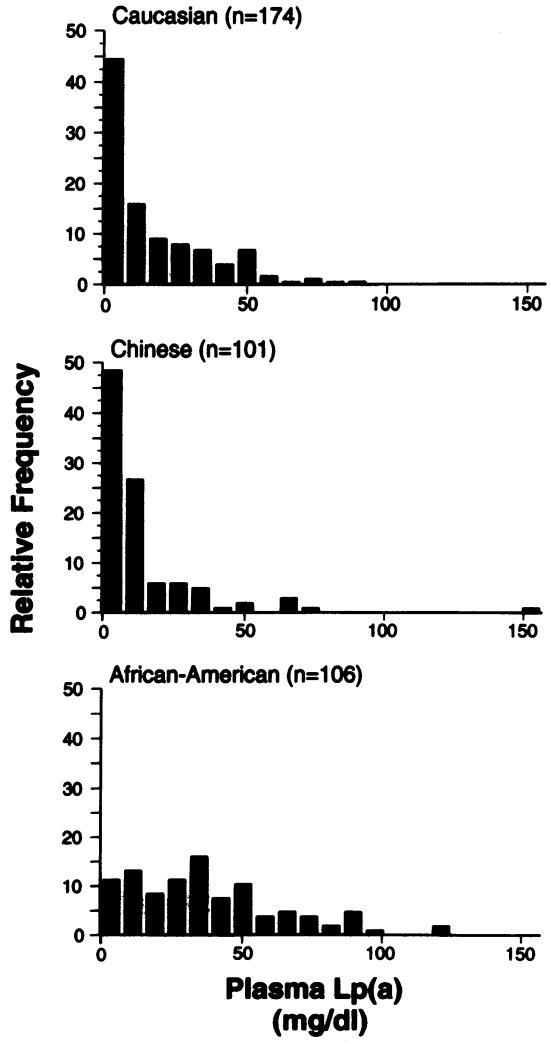

Figure 1. Frequency distributions of plasma Lp(a) concentrations in three populations. Plasma Lp(a) levels were measured in 381 subjects by a sandwich ELISA, as described in Methods. The width of each bar corresponds to a plasma Lp(a) concentration interval of 7.8 mg/dl.

Table II. Description of Plasma Lp(a) Concentrations in U.S. Caucasians, Chinese, and African Americans

Race	Mean	Median	Standard deviation	Skew	Kurtosis	Min.	Max.
	mg/dl	mg/dl	mg/dl			mg/dl	mg/dl
Caucasian	17.9	9.0	19.1	1.33	4.22	<1	88.0
Chinese	14.7	8.0	21.1	3.77	22.27	<1	156.0
African American	37.7	33.0	26.9	0.84	3.33	<1	120.0

In the Caucasians and Chinese, > 45% of the individuals had plasma Lp(a) concentrations less than 9 mg/dl whereas only 12% of the African Americans had plasma Lp(a) levels in this range. In the African Americans, the distribution of plasma Lp(a) concentrations was relatively uniform up to ~ 50 mg/dl, whereas in the other two populations there was a dramatic decrease over this range (Fig. 1).

Since plasma Lp(a) concentrations in the three populations were not normally distributed, the median rather than mean plasma Lp(a) concentrations was used to compare levels. There was no significant difference in median plasma Lp(a) levels in the Caucasian and Chinese populations (Kruskul-Wallis, P = 0.21). In the African Americans, the average plasma concentration of Lp(a) was 2–3-fold higher than in the other two racial groups.

The K-4 encoding region of the apo(a) alleles was analyzed using pulsed-field gel electrophoresis and genomic blotting (12). The alleles were classified according to the estimated total number of K-4 repeats contained within their sequences (*apo(a)12-apo(a)51*). The observed genotype frequencies in each of the racial groups was consistent with Hardy-Weinberg expectations (26). The distributions of apo(a) alleles in the three populations are shown in Fig. 2. In both the Caucasians and African Americans, the distributions of apo(a) alleles were remarkably symmetrical and they were not significantly different from each other (Fisher exact test, P = 0.17). However, the distributions of apo(a) alleles in the Caucasians and African American were significantly different from the Chinese (P < 0.001). In the Chinese population, the mean size of the apo(a) alleles was significantly larger than in the other two groups.

Due to the large number of different sized alleles in these populations, post hoc statistical comparisons of the frequencies of individual alleles must be viewed with caution. Nonetheless, some striking differences were apparent. A single apo(a) allele, *apo(a)-33*, comprised 20% of the alleles in the Chinese sample, whereas the most frequent allele in either the Caucasians or the African Americans contributed only 10% to the alleles in these populations. The high frequency of a single apo(a) allele in the Chinese suggests that the heterozygosity index in this population should be lower. This was indeed the case. The observed frequency of individuals homozygous for apo(a) alleles of identical size was highest in the Chinese (5.5%) and lowest in the Caucasians (2.9%) with the African Americans having an intermediate level (4.7%).

The relationship between apo(a) allele genotypes and plasma concentrations of Lp(a) was examined in each population (Fig. 3). The apo(a) genotypes are represented on the floor (x and y axes) of the 3-D diagram. The vertical bars (z axis) represent the median plasma concentrations of Lp(a) for

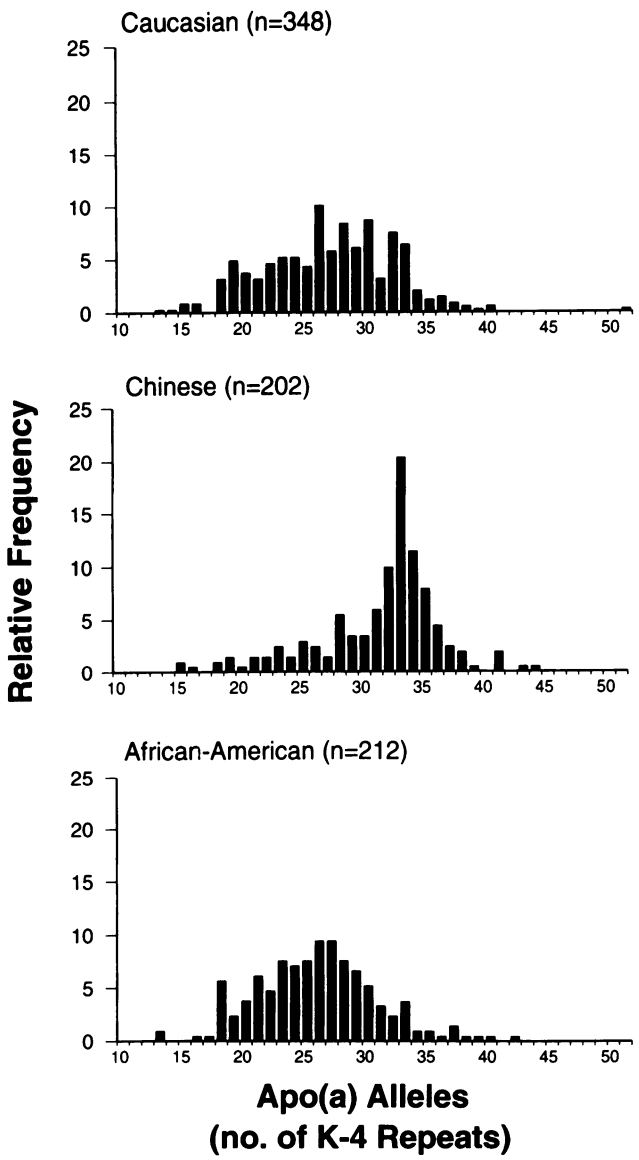

Figure 2. Frequency distributions of apo(a) alleles in three populations. Apo(a) allele size (number of K-4 repeats) was estimated using pulsed-field gel electrophoresis and genomic blotting, as described in the Methods.

all individuals with a given apo(a) genotype. Although the patterns of apo(a) genotypes differed among the three populations, there was a similar inverse relationship between the size of the apo(a) alleles and the plasma concentration of Lp(a), as indicated by the decrease in the height of the bars from left to right. The Spearman rank correlation values for the relationship between the sum of the apo(a) allele sizes (i.e., total number of K-4 repeats) and the plasma levels of Lp(a) were remarkably similar in the three populations: −0.49 (Caucasians and African Americans) and −0.53 (Chinese). The Chinese group had a significantly different "footprint pattern" on the floor of the graph from the other two populations, reflecting the difference in apo(a) allele frequencies (noted above). The Caucasians and African Americans had more similar "footprint patterns", but had distinctly different "skyline patterns," that is, for almost all apo(a) genotypes, the African Americans had higher median plasma Lp(a) levels than the Caucasians. While not apparent from this figure, it should be noted that the variance in plasma Lp(a) levels for each genotype was higher in African Americans (σ^2 = 664) than in the Caucasians (σ^2 = 166) or Chinese (σ^2 = 133).

The higher variance in plasma Lp(a) levels in the African Americans is made more obvious if the total number of K-4 repeats is related to the plasma concentrations of Lp(a) (Fig. 4). In all three populations distribution of plasma Lp(a) levels vs. total number of K-4 repeats was triangular in appearance. The relationship between apo(a) allele size and plasma Lp(a) level is not linear over the entire range. The relationship appears to be biphasic; Caucasian and Chinese individuals with $\gtrsim$ 65 K-4 repeats had uniformly low plasma concentrations of Lp(a) (Fig. 4). The variability in plasma Lp(a) concentrations was greater among individuals with fewer rather than more K4 repeats, and the variability was greatest in the African Americans. This effect may be explained in part by the fact that those individuals with fewer K-4 repeats comprise more than one group of individuals—those with one large and one small apo(a) allele, or those with two medium sized alleles. In contrast, almost all the individuals with a larger total number of K-4 repeats (e.g., > 65) constituted a more uniform group because their K-4 complement could only result from the addition of two large alleles, and not a very small plus a very large allele because such alleles (< 12 or > 51 K-4 repeats) were not observed in this study.

We next examined the relationship between apo(a) allele size and the presence of detectable apo(a) glycoprotein in plasma. The apo(a) isoforms were analyzed by immunoblotting in the entire Chinese group, in all but six of the African Americans, and in 155 of the 174 Caucasians. Immunoblotting data were not obtained in 25 subjects (six African Americans and 19 Caucasians) because blood samples were obtained from these individuals before the development of the more sensitive immunoblotting technique. The alleles were classified into two groups based on whether or not there was clearly detectable plasma apo(a) protein present on immunoblotting. We showed by the analysis of serial dilutions of plasma samples of known Lp(a) concentration that apo(a) alleles associated with a plasma concentration of Lp(a) $\geq$ 0.1 mg/dl were invariably detectable. Only those apo(a) alleles in which no protein product was identified by immunoblotting (i.e., associated with plasma Lp(a) concentrations < 0.1 mg/dl) are refered to as "null" alleles and are included in Fig. 5. Since future refinement of the immunoblotting technique may reveal protein products that are presently undetectable, we have qualified our designation of null alleles by using quotation marks.

The frequency of apo(a) alleles with undetectable protein products was highest in the Caucasians (28.6%), and was similar in the African Americans and Chinese, 14.0% and 13.4%, respectively). The frequency of individuals who had no detectable apo(a) protein by immunoblotting was 5.7% in the Caucasians, 2.0% in the Chinese and 2.8% in the African Americans. In all three populations, the frequency distribution of alleles associated with trace to no detectable apo(a) protein (i.e., < 0.1 mg/dl Lp[a]) was similar to the frequency distribution of alleles associated with detectable apo(a) protein (i.e., $\geq$ 0.1 mg/dl Lp[a]). Thus, despite the overall relationship between apo(a) allele size and plasma Lp(a) level in all three populations, the frequency of low expressing alleles was not skewed towards the larger alleles, as might have been expected based on the overall relationship between apo(a) allele size and plasma level of Lp(a).

Finally, we examined the relationship between apo(a) allele size and plasma level of Lp(a) in the subset of individuals

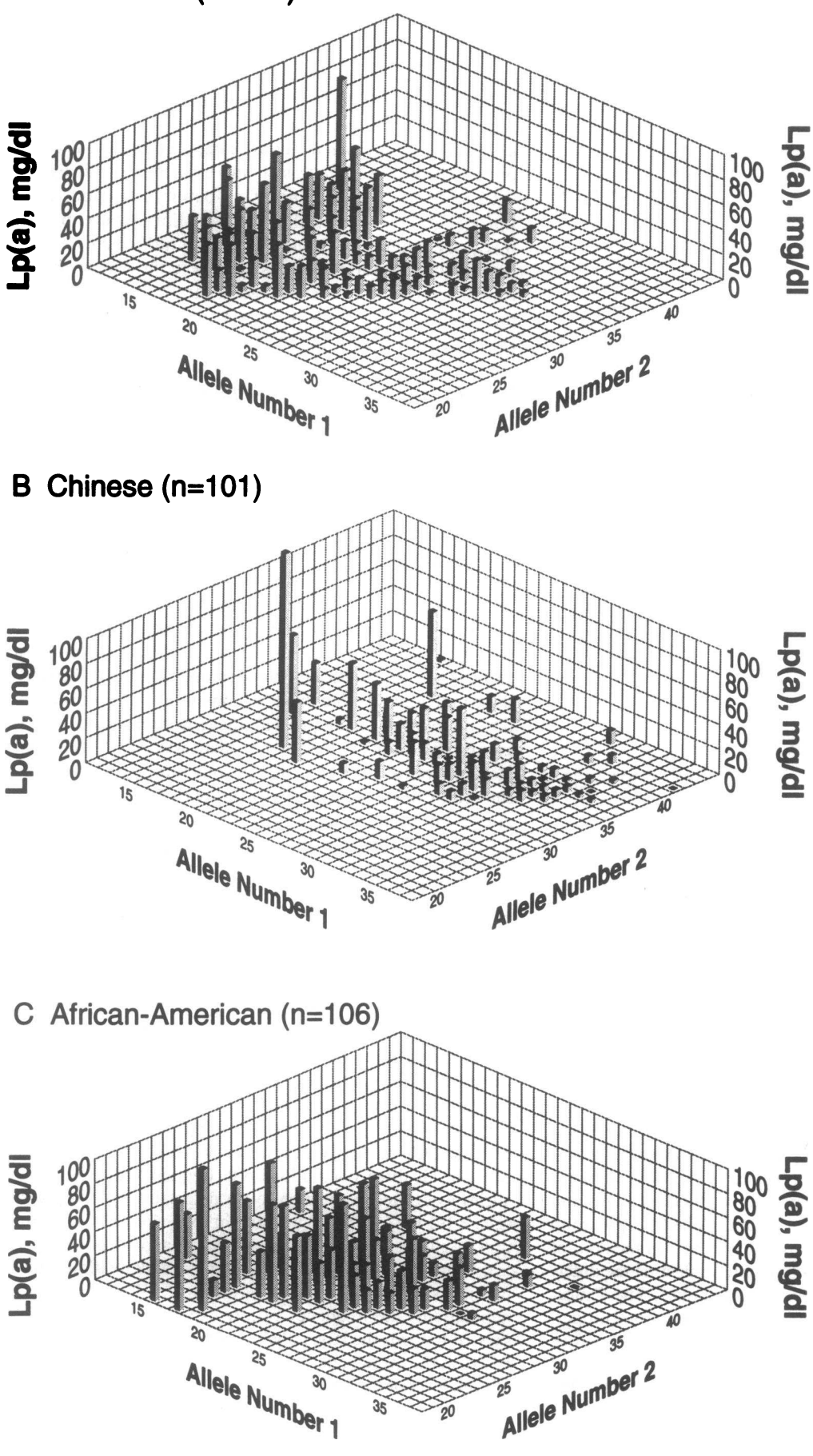

Figure 3. Relationship between apo(a) genotype and plasma Lp(a) concentration in three populations. In each 3-D plot, the number of K-4 repeats of the smaller apo(a) allele (Allele 1) is given on the *x*-axis. The number of K-4 repeats in the larger apo(a) allele (Allele 2) is shown on the *y*-axis. The vertical bars (*z*-axis) represent the median plasma Lp(a) concentration of individuals with each genotype.

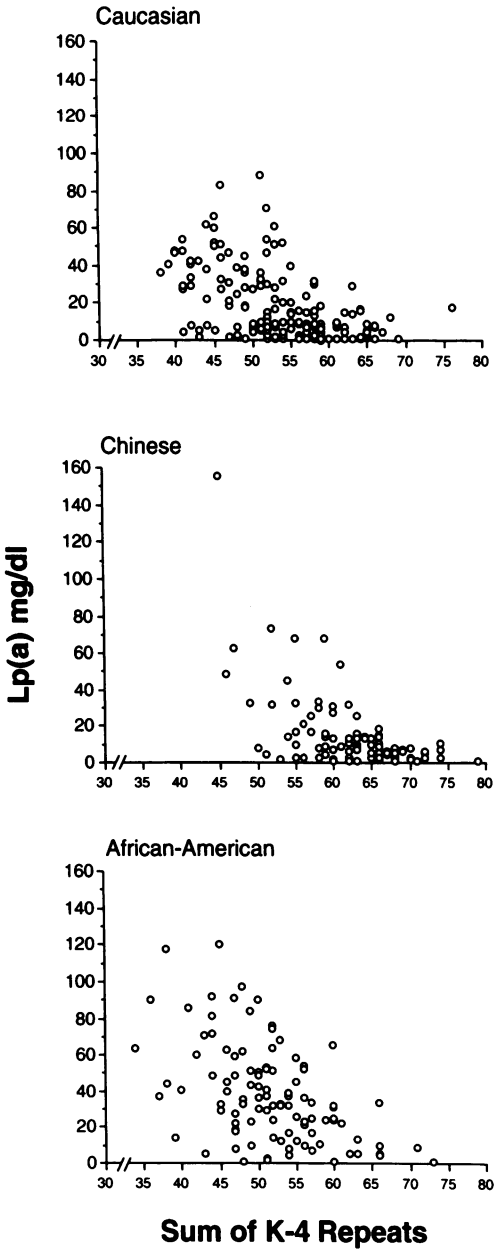

Figure 4. Relationship between the sum of K-4 repeats in the apo(a) alleles and plasma Lp(a) concentration in three populations. The total number of K-4 repeats from each individual, obtained by summing both apo(a) allele sizes, is plotted against the individual plasma Lp(a) concentrations.

who had one allele which expressed no, or only trace, amounts of immunodectable apo(a) protein. For this analysis, we assume that essentially all of the plasma Lp(a) measured by the ELISA is the product of a single expressing allele (Fig. 6). Again, in the Caucasians, there was a distinct triangular relationship between Lp(a) level and the size of the expressing allele. Apo(a) alleles with less than 25 K-4 repeats showed much greater variability in the associated plasma Lp(a) levels. Apo(a) alleles with greater than 28 K-4 repeats consistently were associated with plasma Lp(a) levels less than 10 mg/dl. There were too few individuals with single expressing alleles in the Chinese or African Americans to make any definitive comments about the relationship to plasma Lp(a) level, although the pattern in both groups also appears to be triangular, especially in the African American subset (Fig. 6).

Discussion

This is the first study in which the *apo(a)* gene structure, apo(a) glycoprotein size, and plasma concentration of Lp(a) have been compared in any population. The apo(a) genotype frequencies in all three racial groups were in accordance with Hardy-Weinberg equilibrium, which is consistent with the samples studied being representative of the populations from which they were drawn. This is in contrast to the findings of previous studies that examined the size distribution of apo(a) alleles indirectly by examining the apo(a) isoforms (16, 18). In those studies, not all apo(a) allele protein products were detected, as indicated by the high estimated frequency of null alleles and the large number of subjects with only a single visible isoform. Direct analysis of the *apo(a)* gene obviates the problems associated with nondetection of apo(a) allele protein products when trying to assess the apo(a) genotype frequency.

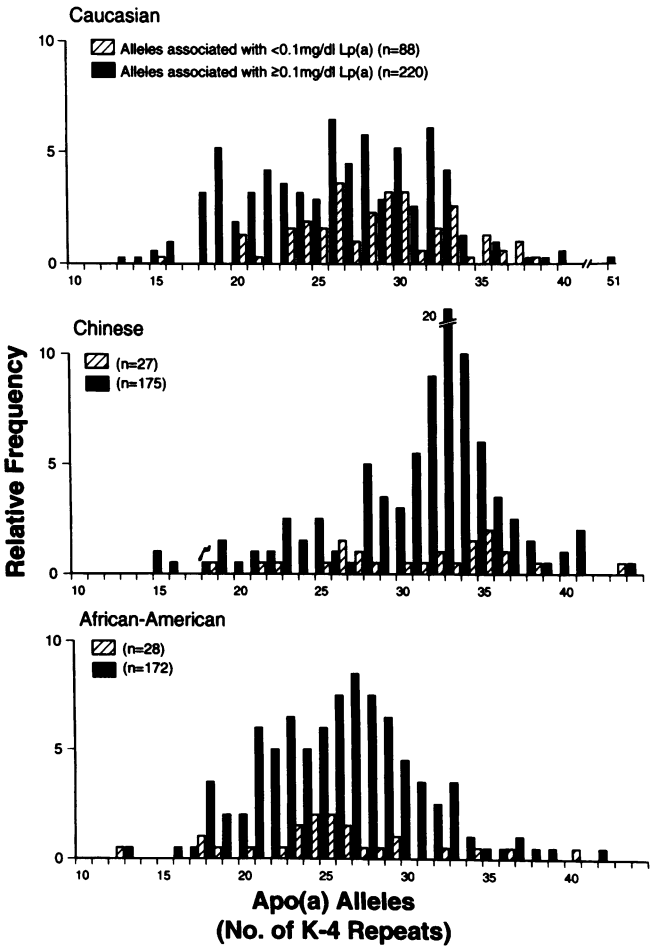

Figure 5. Frequency distributions of expressing and non-expressing apo(a) alleles in three populations. Apo(a) allele and isoform sizes were both determined in each subject. The relative frequencies of apo(a) alleles associated with detectable protein (i.e., ≥ 0.1 mg/dl plasma Lp[a]) and non-detectable protein (i.e., < 0.1 mg/dl plasma Lp[a]) are plotted according to race. In all populations, the non-expressing alleles are spread throughout the entire spectrum of apo(a) allele sizes.

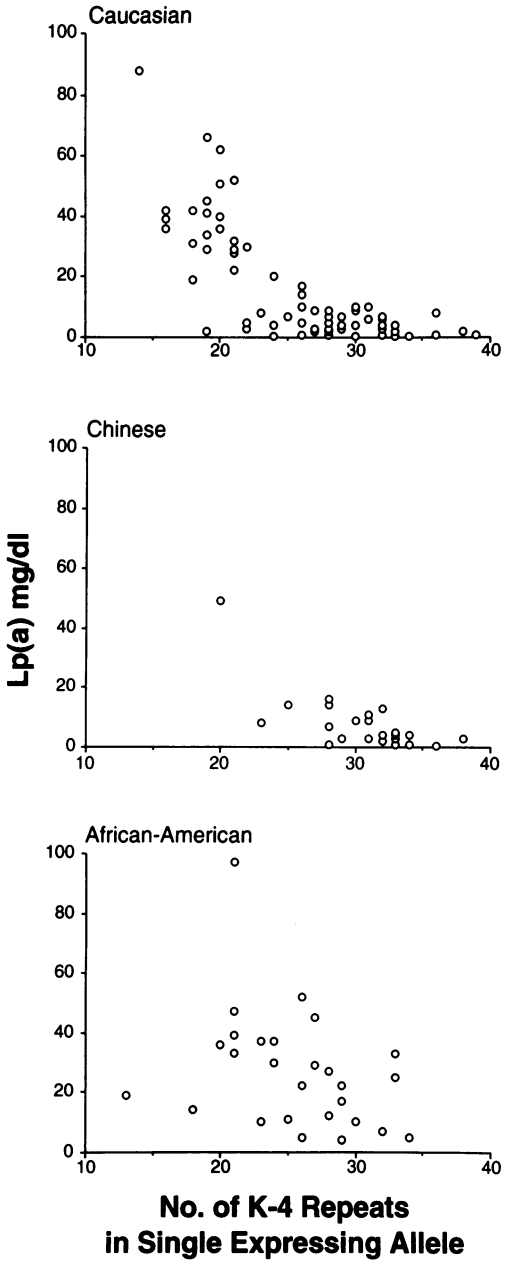

Figure 6. Relationship between the size of a single expressing apo(a) allele and plasma Lp(a) concentration in three populations. A number of individuals in each racial group had only a single expressing apo(a) allele (Caucasians, $n = 68$; Chinese, $n = 23$; African Americans, $n = 27$). The sizes (number of K-4 repeats) of these apo(a) alleles are plotted against the corresponding plasma Lp(a) concentrations revealing a similar pattern for each racial group as that illustrated in Fig. 4.

Consequently, in this study two apo(a) alleles of different sizes were observed in 97% of Caucasians, 95% of Chinese and 94% of African Americans.

A major finding of this study is that the differences in plasma concentrations of Lp(a) in Caucasians, Chinese, and African Americans are not due to in differences in the apo(a) allele size distributions. This conclusion is evident from the following observations: (a) the Caucasians and Chinese had very similar distributions of plasma Lp(a) levels and yet had significantly different apo(a) allele distributions; and (b) the African Americans had a distinctly different plasma Lp(a) distribution from the other two groups, yet their apo(a) allele size distribution was not significantly different from the Caucasians. The similarity in the distribution of apo(a) alleles in Caucasians and African Americans cannot be solely attributed to genetic admixture, since it has been estimated that only ~25% of the genes in African Americans are Caucasian in origin (29). Admixture may contribute to the observed modest skewing of African American plasma Lp(a) concentrations to lower levels (Fig. 1).

It is striking that the relationship between apo(a) allele size and plasma Lp(a) level was so similar in all three ethnic groups. The uniformity of this relationship suggests a fundamental underlying biological mechanism. Interestingly, we found that the relationship between apo(a) allele size and plasma level of Lp(a) was almost identical with previous reports based on the analysis of the apo(a) isoforms (14–16, 18). The reason for this concordance has been elucidated by our simultaneous analyses of the *apo(a)* gene and isoforms. It had been anticipated, given the inverse relationship between apo(a) allele size and plasma Lp(a) level, that the apo(a) alleles associated with no detectable circulating apo(a) would have been of larger size, but this was not the case. The apo(a) alleles which were associated with little to no apo(a) protein had a very similar size distribution to those alleles with detectable protein products. Therefore, the apo(a) alleles undetected by immunoblotting in prior studies (14–16, 18) did not bias the relationship between apo(a) allele size and plasma Lp(a) level.

There are two possible explanations for the fact that apo(a) alleles associated with little or no plasma apo(a) protein are distributed throughout the entire size spectrum. First, the apo(a) alleles associated with very low concentrations of plasma protein may be the result of several different mutations that interfere with apo(a) synthesis or secretion. Given the cysteine-rich nature of this protein, it is likely that even rather conservative amino acid substitutions could disrupt proper folding and subsequent transport of this protein out of the hepatocyte (30). Alternatively, there may be a few common mutations in the *apo(a)* gene which arose early in the gene's evolution and now are associated with apo(a) alleles of widely varying sizes due to subsequent recombination at the locus.

It is interesting that the plasma levels of Lp(a) in the Chinese population were not lower than the Caucasians, given the shift of the apo(a) alleles to the larger sizes (Fig. 2). This is probably due to the fact that a much larger percentage of apo(a) alleles in the Caucasian population were associated with little to no circulating apo(a) protein.

Despite the overall inverse relationship between apo(a) allele sizes and plasma concentrations of Lp(a) in all three populations, there were notable exceptions to this general trend. Apo(a) alleles containing fewer K-4 repeats tend to be associated with higher plasma concentrations of Lp(a) but they can also be associated with little to no detectable plasma apo(a). There is considerably less variation in the plasma levels associated with apo(a) alleles at the high end of the size spectrum. In both the Caucasians and the Chinese, the apo(a) alleles containing more than 28 K-4 repeats were uniformly associated with low plasma levels of Lp(a) (i.e., < 10 mg/dl) (Fig. 6). This is consistent with the observation that apo(a) isoforms of large size are processed inefficiently and secreted at very low

levels in primary cultures of baboon hepatocytes (31). However, it must be noted that in the African Americans, some very large alleles are associated with significant amounts of plasma Lp(a) protein (> 20 mg/dl). Therefore, there is no intrinsic property of large apo(a) isoforms that precludes their synthesis and secretion.

A possible confounding factor in this study is the relative excess of women in the African American sample. In all populations studied to date there has been no significant difference between plasma Lp(a) levels in men and women. The one exception is postmenopausal women where plasma concentrations of Lp(a) are increased by an average of 15% (32). To ensure that the higher plasma Lp(a) levels in the African American population were not due to a relative excess of women in the sample, the median plasma Lp(a) levels were compared in the African American men and women and were found to be not statistically different (30.5 men vs. 33.5 mg/dl women). Similarly, when the sample was stratified to exclude African American postmenopausal women and compared with the total female group of African Americans, the median level was almost identical (32.5 vs. 33.5). Therefore, we are confident that the results of this study are not biased significantly by the differences in sex ratios.

In previous studies, the contributions of apo(a) allele size polymorphism to inter-individual variations in plasma Lp(a) levels have been calculated. However, these calculations are based on the ratio between the variance in apo(a) allele size and the variance in plasma Lp(a) level. As noted previously, for almost every genotype, the plasma concentrations of Lp(a) vary much more in the African Americans than in either of the two other populations. Accordingly, the size polymorphism appears to contribute much less to the inter-individual variability in plasma concentrations of Lp(a) in the African Americans due to the greater variance in plasma Lp(a) levels in this population. Therefore, the significance of such calculations is not clear.

A compelling question that remains unanswered by this study is why African Americans have higher plasma concentrations of Lp(a) than Caucasians despite the similar apo(a) allele distributions. We can definitively state that the higher plasma concentrations of Lp(a) in African Americans are not due to a higher frequency of apo(a) alleles with fewer K-4 repeats. Nor can we explain the higher plasma Lp(a) levels in African Americans by their relatively lower frequency of "null" alleles. Rather, high Lp(a) levels in African Americans appear to be due to a more general mechanism. For almost every apo(a) genotype, the African Americans had higher median plasma levels of Lp(a) than did Caucasians or Chinese. To date, little is known about the genetics of plasma Lp(a) concentrations in African Americans. In Caucasians, the *apo(a)* gene, or sequences linked to it, contribute > 90% to the inter-individual variation in plasma Lp(a) levels (13). The relative contribution of the apo(a) locus to plasma levels of Lp(a) in individuals of African descent has not been determined. The fact that the relationship between apo(a) allele size and plasma Lp(a) level is similar in the African Americans and the Caucasians suggests that sequences at the apo(a) locus contribute importantly to plasma levels of Lp(a) in the African American population. If indeed the *apo(a)* gene is a major determinant of plasma Lp(a) levels in African Americans, there appear to be additional factors, genetic or nongenetic, that modify the effect of the apo(a) allele size in this group.

Acknowledgments

We wish to acknowledge the technical excellence of Kathy Schueler and Tommy Hyatt and the assistance of Carla C. Leffert, R.N.

A. Gaw is the recipient of a British Heart Foundation International Fellowship, and a Travel Scholarship from the James Clerk Maxwell Foundation. E. Boerwinkle is the recipient of a Research Career Development Award and is an Established Investigator of the American Heart Association. H. H. Hobbs is an Established Investigator of the American Heart Association. This work was supported by grants from the Perot Family Fund and National Institutes of Health grants HL-47619 and HL-20948.

References

1. Scanu, A. M. and G. M. Fless. 1990. Lipoprotein(a): heterogeneity and biological relevance. *J. Clin. Invest.* 85:1709–1715.

2. Utermann, G. 1989. The mysteries of lipoprotein(a). *Science (Wash. DC).* 246:904–910.

3. Albers, J. J., S. M. Marcovina, and M. S. Lodge. 1990. The unique lipoprotein(a): properties and immunochemical measurement. *Clin. Chem.* 36:2019–2026.

4. Albers, J. J., J. L. Adolphson, and W. R. Hazzard. 1977. Radio-immunoassay of human plasma Lp(a) lipoprotein. *J. Lipid Res.* 18:331–338.

5. Armstrong, V. W., P. Cremer, E. Eberle, A. Manke, F. Schulze, H. Wieland, H. Kreuzer, and D. Seidel. 1986. The association between serum Lp(a) concentrations and angiographically assessed coronary atherosclerosis: dependence on serum LDL levels. *Atherosclerosis.* 62:249–257.

6. Sandholzer, C., E. Boerwinkle, N. Saha, M. C. Tong, and G. Utermann. 1992. Apolipoprotein(a) phenotypes, Lp(a) concentration and plasma lipid levels in relation to coronary heart disease in a Chinese population: evidence for the role of the apo(a) gene in coronary heart disease. *J. Clin. Invest.* 89:1040–1046.

7. Rhoads, G. G., G. H. Dahlèn, K. Berg, N. E. Morton, and A. L. Dannenberg. 1986. Lp(a) lipoprotein as a risk factor for myocardial infarction. *JAMA (J. Am. Med. Assoc.).* 256:2540–2544.

8. Berg, K., and J. Mohr. 1963. Genetics of the Lp system. *Acta Genet. (Basle).* 13:349–360.

9. Utermann, G., H. J. Menzel, H. G. Kraft, H. C. Duba, H. G. Kemmler, and C. Seitz. 1987. Lp(a) glycoprotein phenotypes: inheritance and relation to Lp(a)-lipoprotein concentrations in plasma. *J. Clin. Invest.* 80:458–465.

10. Lackner, C., E. Boerwinkle, C. C. Leffert, T. Rahmig, and H. H. Hobbs. 1991. Molecular basis of apolipoprotein(a) isoform size heterogeneity as revealed by pulsed-field gel electrophoresis. *J. Clin. Invest.* 87:2077–2086.

11. McLean, J. W., J. E. Tomlinson, W.-J. Kuang, D. L. Eaton, E. Y. Chen, G. M. Fless, A. M. Scanu, and R. M. Lawn. 1987. cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. *Nature (Lond.).* 330:132–137.

12. Lackner, C., J. C. Cohen, and H. H. Hobbs. 1993. Molecular definition of the extreme size polymorphism in apolipoprotein(a). *Hum. Mol. Genet.* 2:933–940.

13. Boerwinkle, E., C. C. Leffert, J. Lin, C. Lackner, G. Chiesa, and H. H. Hobbs. 1992. Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations. *J. Clin. Invest.* 90:52–60.

14. Marcovina, S. M., J. J. Albers, D. R. Jacobs, Jr., L.L. Perkins, C.E. Lewis, B. V. Howard, and P. Savage. 1993 Lipoprotein[a] concentrations and apolipoprotein[a] phenotypes in Caucasians and African Americans. The Cardia Study. *Arterioscler. Thromb.* 13:1037–1045.

15. Marcovina, S. M., Z.H. Zhang, V. P. Gaur, and J. J. Albers. 1993 Identification of 34 apolipoprotein(a) isoforms: differential expression of apolipoprotein(a) alleles between American blacks and whites. *Biochem. Biophys. Res. Commun.* 191:1192–1196.

16. Sandholzer, C., D. M. Hallman, N. Saha, G. Sigurdsson, C. Lackner, A. Csaszar, E. Boerwinkle, and G. Utermann. 1991. Effects of the apolipoprotein(a) size polymorphism on the lipoprotein(a) concentration in 7 ethnic groups. *Hum. Genet.* 86:607–614.

17. Guyton, J. R., G. H. Dahlèn, W. Patsch, J. A. Kautz, and A. M. Gotto,Jr. 1985. Relationship of plasma lipoprotein Lp(a) levels to race and to apolipoprotein B. *Arteriosclerosis.* 5:265–272.

18. Helmhold, M., J. Bigge, R. Muche, J. Mainoo, J. Thiery, D. Seidel, and V. W. Armstrong. 1991. Contribution of the apo(a) phenotype to plasma Lp(a) concentrations shows considerable ethnic variation. *J. Lipid Res.* 32:1919–1928.

19. Parra, H., I. Luyéyé, C. Bouramoué, C. Demarquilly, and J.-C. Fruchart. 1987. Black-white differences in serum Lp(a) lipoprotein levels. *Clin. Chim. Acta.* 167:27–31.

20. Bovet, P., C. Shamlaye, A. Kitua, W. F. Riesen, F. Paccaud, and R. Darioli. 1991. High prevalence of cardiovascular risk factors in the Seychelles (Indian Ocean). *Arterioscler. Thromb.* 11:1730–1736.

21. Menzel, H. J., H. Dieplinger, C. Lackner, F. Hoppichler, J. K. Lloyd, D. R. Muller, C. Labeur, P. J. Talmud, and G. Utermann. 1990. Abetalipoproteinemia with an apoB-100-lipoprotein(a) glycoprotein complex in plasma. *J. Biol. Chem.* 265:981–986.

22. Lipid Research Clinic Program. 1982. Lipid & Lipoprotein Analysis. Manual of Laboratory Operations. Department of Health, Education and Welfare Publ. NIH. Government Printing Office, Washington, D.C., 175–628.

23. Kamboh, M. I., R. E. Ferrell, and B. A. Kottke. 1991. Expressed hypervariable polymorphism of apolipoprotein(a). *Am. J. Hum. Genet.* 49:1063–1074.

24. Nakane, P.K. and A. Kawaoi. 1974. Peroxidase Labeled Antibody. A new method of conjugation. *J. Histochem. Cytochem.* 22:1084–1091.

25. Conover, W.J. 1980. Practical nonparametric statistics. John Wiley & Sons, New York.

26. Guo, W. S., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics.* 48:361–372.

27. Guo, W. S., and E. A. Thompson. 1989. Monte Carlo estimation of p-values for sparse contigency tables. Technical report #187. Department of Statistics. *U of Washington,Seattle, WA.*

28. Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied linear statistical models. *Irwin,Inc.* Homewood, IL.

29. Chakraborty, R., M.I. Kamboh, M. Nwankwo, and R.E. Ferrell. 1992. Caucasian genes in American Blacks: new data. *Am. J. Hum. Genet.* 50: 145–155.

30. Amara, J.F., S.H. Cheng, and A.E. Smith. 1992 Intracellular protein trafficking defects in human disease. *Trend. Cell Biol.* 2:145–149.

31. White, A. L., J. E. Hixson, D. L. Rainwater, and R. E. Lanford. 1994. Molecular basis for "null" lipoprotein (a) phenotypes and the influence of apolipoprotein (a) size on plasma lipoprotein (a) level in the baboon. *J. Biol. Chem.* In press.

32. Heinrich, J., M. Sandkamp, R. Kokott, H. Schulte, and G. Assmann. 1991. Relationship of lipoprotein (a) to variables of coagulation and fibrinolysis in a healthy population. *Clin. Chem.* 37:1950–1954.