JCI The Journal of Clinical Investigation

Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations.

E Boerwinkle, … , G Chiesa, H H Hobbs

J Clin Invest. 1992;90(1):52-60. https://doi.org/10.1172/JCI115855.

Research Article

Plasma lipoprotein(a) [Lp(a)], a low density lipoprotein particle with an attached apolipoprotein(a) [apo(a)], varies widely in concentration between individuals. These concentration differences are heritable and inversely related to the number of kringle 4 repeats in the apo(a) gene. To define the genetic determinants of plasma Lp(a) levels, plasma Lp(a) concentrations and apo(a) genotypes were examined in 48 nuclear Caucasian families. Apo(a) genotypes were determined using a newly developed pulsed-field gel electrophoresis method which distinguished 19 different genotypes at the apo(a) locus. The apo(a) gene itself was found to account for virtually all the genetic variability in plasma Lp(a) levels. This conclusion was reached by analyzing plasma Lp(a) levels in siblings who shared zero, one, or two apo(a) genes that were identical by descent (ibd). Siblings with both apo(a) alleles ibd (n = 72) have strikingly similar plasma Lp(a) levels (r = 0.95), whereas those who shared no apo(a) alleles (n = 52), had dissimilar concentrations (r = -0.23). The apo(a) gene was estimated to be responsible for 91% of the variance of plasma Lp(a) concentration. The number of kringle 4 repeats in the apo(a) gene accounted for 69% of the variation, and yet to be defined cis-acting sequences at the apo(a) locus accounted for the remaining 22% of the inter-individual variation in plasma Lp(a) levels. During the course of these […]

Find the latest version:

https://jci.me/115855/pdf

Apolipoprotein(a) Gene Accounts for Greater Than 90% of the Variation in Plasma Lipoprotein(a) Concentrations

Eric Boerwinkle,* Carla C. Leffert,‡ Jingping Lin,* Carolin Lackner,‡ Giulia Chiesa,‡ and Helen H. Hobbs‡
**Center for Demographic and Populations Genetics, University of Texas Health Science Center in Houston, Houston, Texas 77225; and ‡Departments of Internal Medicine and Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235*

Abstract

Plasma lipoprotein(a) [Lp(a)], a low density lipoprotein particle with an attached apolipoprotein(a) [apo(a)], varies widely in concentration between individuals. These concentration differences are heritable and inversely related to the number of kringle 4 repeats in the apo(a) gene. To define the genetic determinants of plasma Lp(a) levels, plasma Lp(a) concentrations and apo(a) genotypes were examined in 48 nuclear Caucasian families. Apo(a) genotypes were determined using a newly developed pulsed-field gel electrophoresis method which distinguished 19 different genotypes at the apo(a) locus. The apo(a) gene itself was found to account for virtually all the genetic variability in plasma Lp(a) levels. This conclusion was reached by analyzing plasma Lp(a) levels in siblings who shared zero, one, or two apo(a) genes that were identical by descent (ibd). Siblings with both apo(a) alleles ibd (n = 72) have strikingly similar plasma Lp(a) levels (r = 0.95), whereas those who shared no apo(a) alleles (n = 52), had dissimilar concentrations (r = −0.23). The apo(a) gene was estimated to be responsible for 91% of the variance of plasma Lp(a) concentration. The number of kringle 4 repeats in the apo(a) gene accounted for 69% of the variation, and yet to be defined *cis*-acting sequences at the apo(a) locus accounted for the remaining 22% of the inter-individual variation in plasma Lp(a) levels. During the course of these studies we observed the de novo generation of a new apo(a) allele, an event that occurred once in 376 meioses. (*J. Clin. Invest.* 1992. 90:52–60.) Key words: apolipoprotein (a) • lipoprotein (a) • low density lipoprotein

Introduction

Lipoprotein(a) [Lp(a)][1] is a cholesterol ester-rich plasma lipoprotein comprising two attached components: a low density lipoprotein (LDL) particle and a single large glycoprotein, apolipoprotein(a) [apo(a)] (1–3). High plasma levels of Lp(a) are associated with the development of coronary atherosclerosis (4–6) and other vascular diseases (7). The mechanism by which Lp(a) expedites the atherosclerotic process is not known. Apo(a) strongly resembles plasminogen, and it may competitively interfere with plasminogen action in fibrinolysis (8, 9).

Plasma concentrations of Lp(a) vary over a wide range among individuals, but are remarkably stable in any given individual (10). Many physiological, pharmacological, and environmental factors that affect the levels of other plasma lipoproteins have no effect on the plasma concentration of Lp(a) (10). This lack of environmental and physiological influences suggests that plasma Lp(a) levels are largely genetically determined. Consistent with this formulation, early genetic studies suggested that the presence of Lp(a) in plasma was inherited as an autosomal dominant trait (11–13). When more sensitive immunoassays of plasma Lp(a) concentrations were used, it was found that plasma Lp(a) concentrations varied continuously among individuals (14), and the pattern of inheritance indicated that a major gene, as well as polygenic factors, contributed to plasma Lp(a) concentrations (15–17).

Fless et al. (18) and Utermann et al. (19) found that the apo(a) glycoprotein varied in size among individuals. In an important series of studies, Utermann and his colleagues demonstrated that the size of the apo(a) protein is inversely related to the level of plasma Lp(a), thus implicating the apo(a) gene as a major determinant of plasma Lp(a) concentrations (20–23). However, the immunoblotting technique used to type the apo(a) isoforms was not sensitive enough to detect low levels of apo(a) protein, and not all of the apo(a) isoforms were detected. As a result, the frequency distribution of the apo(a) isoforms failed to fit the expectations of Hardy-Weinberg equilibrium (22). In addition, when immunoblotting was employed to examine the segregation of the apo(a) isoforms in families, the results were frequently uninformative, and occasionally inconsistent (24). Further progress required the development of a technique that was more discriminating than immunoblotting in classifying apo(a) alleles.

A potential method to study this polymorphism was suggested by the findings of McLean et al. who discovered that the apo(a) cDNA contains multiple tandem copies of a sequence that encodes a cysteine-rich protein motif called a kringle. The repeated kringle in apo(a) is designated kringle 4 because it closely resembles the fourth kringle in plasminogen. McLean et al. proposed that the apo(a) isoforms are of different size because of variations in the numbers of kringle 4-encoding repeats in the apo(a) gene (25). This hypothesis was supported by studies of the apo(a) mRNA and gene structure (26–28). In attempt to devise a way to measure the size of the apo(a) gene in different individuals, we previously identified a large restriction fragment from the apo(a) gene which contains most, if not all, of the kringle 4–encoding sequences (29). The size of this fragment was too large to be examined by standard electrophoresis techniques. Accordingly, we used pulsed-field gel electrophoresis to size this large restriction fragment and 19 fragments of different length were identified. A total of 103 unrelated

Address reprint requests to Dr. Hobbs, Department of Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75235.
Received for publication 6 December 1991.

1. *Abbreviation used in this paper:* Lp(a), lipoprotein (a).

J. Clin. Invest.
© The American Society for Clinical Investigation, Inc.
0021-9738/92/07/0052/09 $2.00
Volume 90, July 1992, 52–60

Caucasians were evaluated and 94% were heterozygous for fragments of two different sizes. The length polymorphism was used as a genetic marker to analyze the segregation of the apo(a) gene in 12 Caucasian families. It was found that within a given family, sibling pairs with identical apo(a) genotypes tended to have very similar plasma Lp(a) levels (29). However, individuals with the same apo(a) genotypes who were members of different families often had significantly different plasma concentrations of Lp(a). Taken together, these observations suggest that the apo(a) gene is the major determinant of plasma Lp(a) levels and that cis-acting DNA sequences at or near the apo(a) locus, other than the number of kringle 4 repeats, contribute importantly to plasma Lp(a) concentrations.

In the current study, we analyzed the segregation of the apo(a) gene and Lp(a) levels in 48 Caucasian pedigrees to determine the contribution of the apo(a) gene (or closely linked loci) to the plasma concentrations of Lp(a). The genetic architecture (30) of plasma Lp(a) concentrations was defined at three levels: the polygenic heritability, the total genetic contribution of the apo(a) gene, and the effects of length variation in the apo(a) gene. In addition, we describe the de novo generation of a new apo(a) allele of different size within a family.

Methods

Subjects. Plasma Lp(a) concentrations were measured in a sample of 288 fasting individuals from 48 Caucasian American families living in the greater Dallas, Texas area. Families in which both parents and at least three children were available for sampling were selected for study. None of the families had evidence of a monogenic hyperlipidemia. In one family, F153, several members (who are denoted in Table II) had very low LDL-cholesterol levels, suggesting the possible existence of familial hypobetalipoproteinemia. For the random effects analysis of variance (see below), the families were augmented with a sample of 107 unrelated individuals. Preliminary findings on a subset of these unrelated individuals have been reported previously (29).

Phlebotomy was carried out after an overnight fast. A total of 30 ml of blood was collected from each individual in vacutainer tubes containing sodium-EDTA. The plasma was separated within one hour of collection by centrifugation at 2,000 *g* for 15 min at 4°C. Multiple 50 μl aliquots of plasma were stored at −70°C and Lp(a) levels were assayed within 4 wks.

Pulsed-field gel analysis of the apo(a) gene. A total of 15 ml of blood was maintained at room temperature prior to transfer to two LeucoPREP tubes (Becton, Dickinson & Co., Lincoln Park, NJ). Lymphocytes were isolated and embedded in agarose plugs as previously described (29). The agarose-cellular plugs were incubated twice with 40 U of KpnI in 170 μl of the buffer suggested by the manufacturer (New England Biolabs, Beverly, MA). The digested cellular-agarose plugs were subjected to pulsed-field gel electrophoresis in a vertical submarine gel apparatus with a transverse alternating field (Geneline I, Beckman Instruments, Inc., Fullerton, CA) using low-endosmosis coefficient agarose, TAFE buffer, and λ phage concatamer standards (Beckman Instruments, Inc.) as described by Lackner et al. (29). After 18 hours of electrophoresis, the gel was stained with ethidium bromide and photographed. The DNA was transferred and fixed to nylon membrane (Biotrans, ICN Biomedicals, Costa Mesa, CA). MP-1, a 1.5-kb PstI genomic fragment from the kringle 4-encoding region of the apo(a) gene (29) which had been subcloned into M13mp18, was used to generate a ^{32}P-radiolabeled single-stranded probe (31). The filter was incubated overnight at 42°C in hybridization solution containing 5 × 10^6 cpm/ml of the single-stranded apo(a)-specific probe. Hybridizations were carried out in a rotating incubator (model 310, Robbins Scientific Corp., Sunnyvale, CA). Filters were washed as described by Lackner et al. (29) and exposed to film.

Immunoblotting of plasma apo(a). An aliquot of frozen plasma (between 1 and 30 μl) containing 1 μg of Lp(a) was brought up to a total volume of 30 μl using phosphate-buffered saline. The sample was mixed with 20 μl of buffer A which contained 15% filtered SDS (wt/vol), 8 M urea, 5 mM dithiothreitol, and 62.5 mM Tris at pH 7.5 and with 50 μl buffer B (10% glycerol [vol/vol], 2.3% SDS [wt/vol], 0.025% bromophenol blue [wt/vol], 5% β-mercaptoethanol [vol/vol], and 50.0 mM Tris at pH 6.8). The samples were boiled for 10 min before loading onto a 3–7% gradient polyacrylamide gel with SDS. A total of 1 μg of purified LDL (molecular weight of apo B is ~ 513 kD) was used as a size standard. The electrophoresis, transfer to nitrocellulose, and hybridization conditions were exactly as preciously described except that IgG-1A^2, the apo(a)-specific antibody, was radiolabeled directly with ^{125}I to a specific activity of 5 × 10^6 cpm/ml (29). The filters were washed, dried and exposed to XAR-5 film (Eastman Kodak Co., Rochester, NY) at −70°C with an intensifying (Lightening Plus, Dupont Co., Wilmington, DE).

Plasma lipid and lipoprotein assays. Measurement of plasma Lp(a) concentrations were performed at GeneScreen, Dallas, TX, using a sensitive enzyme-linked immunosorbent sandwich assay (ELISA), as described (32). In this assay, Lp(a) was captured by a polyclonal rabbit anti-human Lp(a) antibody and then detected by a monoclonal anti-human Lp(a) antibody, IgG-1A^2. Plasma Lp(a) standards were obtained from Immuno, Vienna, Austria. Total cholesterol and triglyceride levels were measured enzymatically using commercially available kits (Boehringer Mannheim, Indianapolis, IN; Sigma Chemical Co., St. Louis, MO). Plasma lipoproteins were quantified in the laboratory of Dr. Scott Grundy (University of Texas Southwestern Medical Center) according to the procedures of the Lipid Research Clinic (33).

Statistical methods. The distribution of plasma Lp(a) concentration was positively skewed in these data, and thus all analyses were carried out both on the raw and square-root transformed data. For each analysis, the primary inferences were identical whether the raw or transformed data was used.

The contribution of unmeasured polygenic variation to the inter-individual variability of plasma Lp(a) concentrations ($\sigma_{Lp(a)}{}^2$) was assessed from the extent of familial aggregation of Lp(a) levels in the sample of pedigrees. The ratio of the polygenic variance component ($\sigma_{pg}{}^2$) to $\sigma_{Lp(a)}{}^2$ was estimated by maximum likelihood principles as implemented in the computer program PAP V3.0 (34).

Sibling-pair linkage methods were used to estimate the overall contribution of genetic variation in and around the apo(a) gene ($\sigma_{apo(a)}{}^2$) to plasma Lp(a) levels (35, 36). These methods are most frequently employed to detect linkage between a marker and a quantitative trait locus, but can also be used to define the overall contribution of a candidate gene to a quantitative phenotype. For each sibling pair, three new variables were considered: y_j, the squared difference of plasma Lp(a) concentrations in sibship j, f_{1j}, an indicator variable describing whether or not the jth sib pair shares only 1 allele identical by descent (ibd), and π_j, the proportion of alleles ibd in sibship j. π_j can take on the values 0, ½, or 1. $E(y_j)$ is the expected value of an individual's Lp(a) concentration. Assuming there is no recombination as would be the case for a candidate gene, Haseman and Elston (30) show that:

$$E(y_j) = \alpha + \beta\pi_j + \gamma f_{1j} \quad (1)$$

where

$$\alpha = 2\sigma_{apo(a)}{}^2 + \sigma_e{}^2 \quad (2)$$

$$\beta = -2\sigma_{apo(a)}2 \quad (3)$$

$$\gamma = -\sigma_{d.}{}^2 \quad (4)$$

In Eqs. 2–4, $\sigma_e{}^2$ is a residual variance component describing the effects of factors other than the apo(a) gene on Lp(a) levels, and $\sigma_d{}^2$ describes the dominance effects at the apo(a) locus on Lp(a) levels. An estimate of the overall contribution of the apo(a) gene to plasma Lp(a) concentrations can be made by examining the regression of the squared differ-

Table I. Correlations of Plasma Lp(a) Concentrations between Family Members

	n	Lp(a)	$\sqrt{\text{Lp(a)}}$
Spouses	48	0.17 [−0.12, 0.43]*	0.17 [−0.12, 0.43]
Parent-offspring	400	0.44§ [0.36, 0.52]	0.48§ [0.40, 0.55]
Midparent-offspring	200	0.59§ [0.49, 0.67]	0.61§ [0.51, 0.69]
Siblings (all)	284	0.28§ [0.16, 0.39]	0.32§ [0.21, 0.43]
Siblings sharing no alleles ibd	52	−0.23 [−0.47, 0.05]	−0.25 [−0.48, 0.02]
Siblings sharing one allele ibd	159	0.15 [−0.16, 0.30]	0.19‡ [0.04, 0.34]
Siblings sharing two alleles ibd	73	0.95§ [0.92, 0.97]	0.96§ [0.94, 0.97]

* 95% confidence interval. ‡ $P < 0.05$. § $P < 0.001$.

ence between the Lp(a) levels of siblings who share none, one, or all apo(a) alleles ibd. The regression analyses were performed both unweighted and weighted, as suggested by Amos et al. (36), with nearly identical results. Therefore, only the results of the unweighted analyses are presented. Even though the sibships were typically larger then size two, the above method has been shown to be valid when overlapping sibling pairs are analyzed as though they were independent (36).

The contribution of length variation in the apo(a) gene, as measured by pulsed-field gel electrophoresis, to Lp(a) concentrations, σ_{length}^2, was estimated using a random effects analysis of variance (37). A random effects or type II model was selected because of the large number of potential genotypes at the apo(a) locus (38).

Results

Plasma Lp(a) concentrations were measured in 288 individuals from 48 pedigrees. There was no significant effect of age, sex, or the concentration of other plasma lipoproteins on the plasma level of Lp(a), so these factors were not considered further in the family members (data not shown). There were significant correlations between the plasma Lp(a) levels of parents and offspring ($r = 0.44$), and siblings ($r = 0.28$), but not between spouses ($r = 0.17$) (Table 1). By using standard biometrical genetic analyses, it was estimated that 85% (±8%) of the inter-individual variance of Lp(a) concentrations was attributable to polygenic effects ($\sigma_{pg}^2/\sigma_{Lp(a)}^2$) (or 88% (±6.5) when the square-root of the plasma Lp(a) levels was used).

Pulsed-field gel electrophoresis and genomic blotting of KpnI digested-genomic DNA was performed to assess the size of the kringle 4–encoding region of the apo(a) alleles in each family member. 16 of the 19 previously described apo(a) alleles were observed in the sample, and their frequencies did not differ significantly from those previously described from the same population (29). In general, there was an inverse relationship between the size of the apo(a) allele and the plasma level of Lp(a). One way to illustrate this phenomenon is to examine the relationship between the plasma Lp(a) concentrations and the apo(a) allele size in the individuals who had one of the two most common alleles, *apo(a)14* or *apo(a)15* plus a different allele (Fig. 1). Individuals with one copy of *apo(a)14* or *apo(a)15* plus one copy of *apo(a)2-apo(a)4* tended to have high plasma Lp(a) levels (> 30 mg/dl). If the second allele was *apo(a)5-apo(a)7*, the Lp(a) levels were lower (15–30 mg/dl). If the second allele was larger than *apo(a)8*, the plasma concentrations of Lp(a) were low [< 10 mg/dl, excluding *apo(a)10*].

In the population as a whole, different apo(a) genotypes, as determined by pulsed-field gel electrophoresis, were associated with significantly different plasma levels of Lp(a) ($P < 0.001$ for both the raw and transformed data). A random effects analysis of variance (37) was used to determine the contribution of the length variation in the apo(a) gene to the distribution of plasma Lp(a) in 203 unrelated Caucasians. For the raw data, 69% of the variation in Lp(a) concentrations was attributable to inter-individual differences in the number of kringle-4 repeats. The square-root transformation had little effect on this value (66% vs. 69%).

Although length variation in the apo(a) gene had a profound influence on Lp(a) concentrations, there were several exceptions to the general trend. Fig. 2 shows two pedigrees in which an apo(a) allele of the same size, *apo(a)6*, segregates. In the two pedigrees this allele gives rise to very different plasma concentrations of Lp(a). In *A*, the *apo(a)6* allele of the father (*a*), is inherited by three of his offspring (*c*, *d*, and *f*). The father, as well as the three offspring, have modest plasma Lp(a) concentrations (6 mg/dl, and 7, 5, and 3 mg/dl, respectively). In the family shown in *B*, individual h, who is also heterozygous for an allele the size of *apo(a)6*, and has a high plasma Lp(a) concentration (51 mg/dl). Of her four children, only the second child (*j*) inherited *apo(a)6* and she is the only offspring with a comparable plasma level of Lp(a) (51 mg/dl). Therefore, in these two families, the same sized apo(a) allele (*apo(a)6*) segregated with very different plasma levels of Lp(a). This was true even though the other alleles at the apo(a) locus in the families were similar (*apo(a)13-apo(a)17*). These findings suggest that factors at the apo(a)

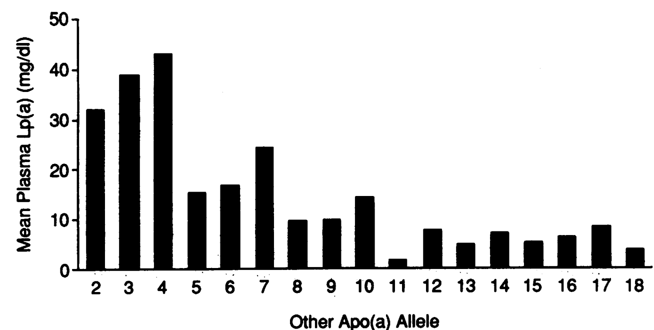

Figure 1. Lp(a) levels in individuals heterozygous for apo(a)14 or apo(a)15 allele. In the sample of 288 family members and 107 unrelated individuals, there were 194 individuals with either *apo(a)14* or *apo(a)15*. The average Lp(a) levels (y-axis) for individuals with each genotype are plotted against the other apo(a) allele (x-axis).

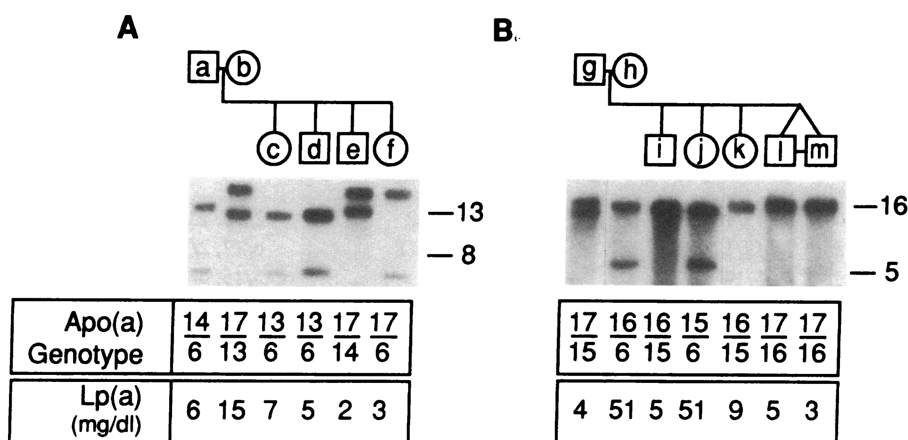

Figure 2. Genomic blot of the apo(a) gene from two unrelated families with *apo(a)6*. High molecular weight leukocyte DNA from members of two unrelated families was digested with KpnI, size-fractionated on a pulsed-field gel, transferred to a nylon membrane, and hybridized with a single-stranded apo(a)-specific probe (MP-1) as described in the Methods. The filter was exposed to Kodak XAR-5 film for 18 h with an intensifying screen. The plasma concentrations of Lp(a) were measured using an ELISA assay as described in the Methods. The *apo(a)6* allele segregates with a low (*A*) and high (*B*) plasma concentration of Lp(a) in two different pedigrees.

locus, in addition to the number of kringle 4 repeats, strongly influence the plasma Lp(a) concentration.

Another instance in which apo(a) alleles of the same size are associated with different amounts of circulating apo(a) protein, is shown in Fig. 3. In this family, the mother (b) is homozygous for *apo(a)12,* so all of the children (*c–f*) are heterozygous for that allele. Based on the genomic blot, it cannot be determined which of the two *apo(a)12* alleles each child inherited from their mother. However, analysis of the apo(a) protein isoforms reveals that three of the offspring (*c, d,* and *f*) have no detectable apo(a) protein corresponding to *apo(a)12.* Only offspring *e* has a band the same size as the isoform of the mother. This suggests that the mother is heterozygous and has one *apo(a)12* that produces no detectable circulating apo(a) protein which she gave to *c, d,* and *f* and another that is associated with the production of a moderate amount of apo(a) protein which she donated to offspring *e*.

To confirm that cis-acting sequences at the apo(a) locus are

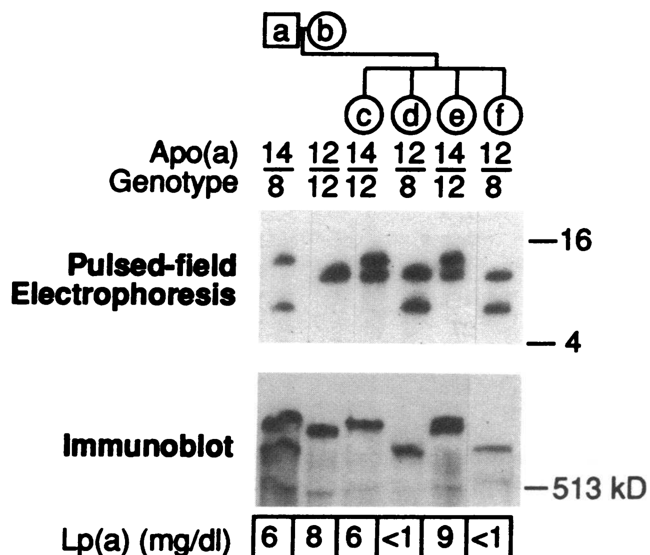

Figure 3. Genomic blot of apo(a) gene and immunoblot of apo(a) protein in a pedigree. The plasma Lp(a) concentrations were measured and the genomic blot and immunoblot was performed as described in Fig. 2 and in the Methods.

responsible for the observed differences in plasma Lp(a) concentrations in individuals with apo(a) alleles of the same size, the plasma Lp(a) concentrations were compared in sibling pairs who shared all, one, or no apo(a) alleles ibd. In 40 of the 48 families, all four parental apo(a) alleles could be differentiated using pulsed-field gel electrophoresis. In six families (including the one shown in Fig. 3), the length polymorphism was uninformative because one of the parents was apparently homozygous for the same sized apo(a) allele, and in two families one parent was not available for sampling; these eight families were not included in the sibling-pair analysis. The families were selected at random, so in some families all the plasma concentrations of Lp(a) were low (i.e., < 5 mg/dl) reflecting the highly skewed distribution of plasma Lp(a) levels in the Caucasian population.

In the 40 families in which the segregation of each parental allele could be distinguished, 72 sibling pairs shared both, 52 shared none, and 159 pairs shared one parental allele ibd. The apo(a) genotypes and plasma Lp(a) levels of these sibling pairs are given in Table II and can be compared to those of the other siblings and parents. The sibling pairs who had plasma levels of Lp(a) that were similar to each other and significantly different from the other siblings are denoted by an asterisk. In 24 families, at least one sibling pair had inherited identical apo(a) alleles and one sibling pair had no apo(a) alleles ibd (Table III). In 21 of the 24 sibling pairs (denoted by asterisks), the mean difference between plasma Lp(a) levels in the sibling pair who shared no apo(a) alleles ibd was twice that of the sibling pair who shared both apo(a) alleles ibd.

Fig. 4 shows the scatter plot of Lp(a) levels of siblings pairs that share (*A*) both or (*B*) no apo(a) alleles ibd. Lp(a) levels for the older (sibling 1) and younger (sibling 2) sibling are plotted on the horizontal and vertical axis, respectively. The correlation coefficient for Lp(a) levels between siblings who share both apo(a) alleles ibd was very high ($r = 0.95$), whereas there was a negative correlation ($r = -0.23$) between the Lp(a) concentration of siblings who share no apo(a) alleles ibd. Similar results were obtained for the square-root transformed data (Table I). Owing to the highly skewed distribution of Lp(a) in the population, many of the sibling pairs had very low Lp(a) levels. Therefore, the same comparison was made in the sibling pairs with apo(a) alleles ibd ($n = 48$) who had plasma Lp(a) levels over 5 mg/dl and the correlation coefficient remained very high ($r = 0.94$).

Table II. Apo(a) Genotypes and Lp(a) Levels in Sibling Pairs with Apo(a) Alleles ibd and Their Family Members

	Family no.	Sib with identical apo(a) genotypes: Apo(a) genotype	Plasma Lp(a)			Other Sibs: Apo(a) genotypes	Plasma Lp(a)	Parents (father, mother): Apo(a) genotype	Lp(a)
				mg/dl			*mg/dl*		*mg/dl*
1.	F141	2/13	26		41	13/16, 15/16, 2/15	1, 6, 32	2/16, 13/15	36, 6
2.	F168	2/17	*40	*42	*55	12/17, 12/17	< 1, < 1	6/17, 2/12	40, 54
3.	F162	4/5	*44	*47	*58	5/14	21	5/15, 4/14	38, 19
4.	F24	4/10	*41		*42	10/14, 6/14	< 1, 16	4/14, 6/10	31, 62
5.	F154	4/10	58		34	12/15, 4/15, 4/15	14, 32, 43	10/15, 4/12	32, 52
6.	F161	4/13	*49		*52	13/13	10	4/13, 13/14	50, 9
7.	F158	4/14	*36		*48	8/14	10	4/8, 14/15	48, 7
8.	F156	4/15	*47		*49	10/15	7	14/15, 4/10	4, 42
9.	F135	4/15	*55	*56	*66	14/15, 14/15	3, 3	14/15, 4/14	4, 45
10.	F154	4/15	32		43	12/15, 4/10, 4/10	14, 34, 58	10/15, 4/12	32, 52
11.	F142	5/9	*49		*56	9/16	4	5/16, 9/10	72, 21
12.	F145	5/12	64		98	11/16, 11/16, 12/16, 5/11	< 1, 3, 8, 51	5/16, 11/12	47, 5
13.	F129	6/6	7		7	6/16, 6/16, 15/16	5, 6, 6	6/15, 6/16	4, 15
14.	F149	6/13	5		7	14/17, 6/17	2, 3	6/14, 13/17	6, 15
15.	F129	6/16	5		6	15/16, 6/6, 6/6	6, 7, 7	6/15, 6/16	4, 15
16.	F146	7/11	*28		*44	9/17, 9/11, 9/11	< 1, 5, 6	7/9, 11/17	22, 9
17.	F166	8/12	5		7	12/13, 13/15	17, 19	8/13, 12/15	9, 10
18.	F150	8/15	4	6	8	7/15	5	15/16, 7/8	5, 5
19.	F146	9/11	5		6	9/17, 7/11, 7/11	< 1, 28, 44	7/9, 11/17	22, 9
20.	F124	9/16	*7		*9	4/16, 5/9, 4/5	22, 50, 75	5/16, 4/9	54, 27
21.	F134	9/17	6	8	11	16/17, 9/18	12, 12	9/16, 17/18	15, 3
22.	F137	10/13	< 1		3	13/14, 13/14, 13/14	< 1, 1, 1	10/14, 13/15	< 1, 2
23.	F21	11/15	< 1		< 1	15/17, 8/11	< 1, 7	8/15, 11/17	5, < 1
24.	F143	11/15	1		2	14/15, 14/15, 14/15, 7/11	< 1, < 1, < 1, 17	7/15, 11/14	28, < 1
25.	F164	11/15	1	1	1	15/16	3	11/16, 8/15	2, 4
26.	F145	11/16	*< 1		*3	12/16, 5/11, 5/12, 5/12	8, 51, 64, 98	5/16, 11/12	47, 5
27.	F167	12/13	*15	*15	*21	12/17, 13/13	1, 5	13/17, 12/13	12, 1
28.	F138	12/14	1		5	13/15	1	14/15, 12/13	4, 2
29.	F152	12/14	*< 1		*< 1	8/13	16	12/13, 8/14	< 1, 30
30.	F160	12/14	*3		*6	10/15, 10/14	21, 34	10/12, 14/15	27, 10
31.	F125	12/15	1		1	14/15, 7/12	< 1, 28	12/14, 7/15	< 1, 32
32.	F168	12/17	*< 1		*< 1	2/17, 2/17, 2/17	40, 42, 55	6/17, 2/12	40, 54
33.	F126	12/18	< 1		< 1	15/18	< 1	12/15, 9/18	3, < 1
34.	F137	13/14	1	1	< 1	10/13, 10/13	< 1, 3	10/14, 13/15	< 1, 2
35.	F136	13/15	4		4	16/17, 15/17, 8/16	< 1, 3, 3	13/17, 15/16	1, 9
36.	F135	14/15	*3		*3	4/15, 4/15, 4/15	56, 56, 66	14/15, 4/14	4, 45
37.	F143	14/15	< 1	< 1	< 1	11/15, 11/15, 7/11	1, 2, 17	7/15, 11/14	28, < 1
38.	F131	14/15	< 1	< 1	4	15/18	< 1	14/18, 5/15	< 1, < 1
39.	F132	14/17	1	1	3	14/15, 15/17	2, 3	14/17, 14/15	8, < 1
40.	F165	14/18	8		8	15/18, 15/18	< 1, 9	14/15, 10/18	16, < 1
41.	F157	15/16	5		9	16/17, 16/17, 6/15	3, 5, 51	15/17, 6/16	4, 51
42.	F159	15/17	1	5	2	10/17	< 1	16/17, 10/15	< 1, 6
43.	F153	16/16	*1‡	*4‡	*5‡	7/16	27	7/16, 12/16	52, 1
44.	F157	16/17	3		5	15/16, 15/16, 6/15	5, 9, 51	15/17, 6/16	4, 51
45.	F165	15/18	< 1		9	14/18, 14/18	8, 8	14/15, 10/18	16, 1

* Sibling pairs with identical apo(a) genotypes who have Lp(a) levels which are significantly different from all the other siblings. ‡ These individuals have a plasma LDL-cholesterol concentration less than the 5th percentile when compared to age and sex-matched controls.

The overall contribution of the apo(a) gene to plasma Lp(a) concentrations was estimated by examining the regression of the squared difference of Lp(a) levels between siblings (y_j) based on the proportion of apo(a) alleles shared ibd (π_j). The dominance deviations at the apo(a) locus (γ) was not significantly different from zero ($\gamma = 1.27$ for untransformed Lp(a) levels) so was not considered in further analyses. The simple linear regression of the squared difference of Lp(a) levels between siblings on the proportion of apo(a) alleles shared ibd is graphically presented in Fig. 5. The average squared dif-

Table III. Lp(a) Levels in Sibling Pairs in the Same Family Who Share Both or No Apo(a) Alleles Identical by Descent

	Siblings sharing both apo(a) alleles			Siblings sharing no apo(a) alleles			
		Lp(a)		Genotype Lp(a)			
Family	Genotype	Sib1	Sib2	Sib1		Sib2	
141*	2/13	41	26	13/16	1	2/15	32
24*	4/10	42	41	4/10	42	6/14	16
154	4/10	58	34	12/15	14	4/10	34
154*	4/15	32	43	12/15	14	4/10	58
124*	4/16	22	36	5/9	50	4/16	22
145*	5/12	98	64	11/16	1	5/12	98
149*	6/13	7	5	6/13	7	14/17	2
146*	7/11	44	28	7/11	44	9/17	1
166*	8/12	5	7	8/12	5	13/15	19
146*	9/11	5	6	9/17	1	7/11	28
124*	9/16	9	7	9/16	9	4/5	75
134	9/17	11	8	9/18	12	16/17	12
21*	11/15	1	1	15/17	1	8/11	7
143*	11/15	2	1	7/11	17	14/15	1
145*	11/16	1	3	12/16	8	5/11	51
138	12/14	1	5	13/15	1	12/14	1
160*	12/14	3	6	10/15	21	12/14	3
152*	12/14	1	1	12/14	1	8/13	16
125*	12/15	1	1	7/12	28	14/15	1
136*	13/15	4	4	13/15	4	8/16	3
143*	14/15	1	1	7/11	17	14/15	1
143*	14/15	1	1	7/11	17	14/15	1
157*	15/16	5	9	6/15	51	16/17	5
157*	16/17	5	3	6/15	51	16/17	3

* The difference in Lp(a) levels in the siblings sharing no apo(a) alleles is at least twice that of the siblings sharing both alleles.

ferences are 1248, 654, and 58 $(\mathrm{mg/dl})^2$ for those sibling pairs that share no, one, and both of their apo(a) alleles ibd, respectively. There was a heteroscedastic distribution of squared Lp(a) differences among the three groups, so weighted regression analysis was performed, as suggested by Amos et al. (36); the results were similar for the weighted and unweighted analyses (data not shown). The linear regression line that best fits these data was equal to $y_j = 1249.2 - 1190.6\pi_j$. These parameter estimates combined with algebraic manipulation of Eqs. 2 and 3 yield estimates of $\sigma_{apo(a)}^2$ and the residual variance component, σ_e^2. For the raw untransformed data, $\sigma_{apo(a)}^2$ and σ_e^2 were equal to 595.3 and 58.6, respectively. As a ratio, these results indicate that 91% of the variation of plasma Lp(a) concentrations among individuals was attributable to genetic variation in the apo(a) gene ($\sigma_{apo(a)}^2/\sigma_{apo(a)}^2 + \sigma_e^2$). For the square-root transformed data these values were 7.00%, 0.86%, and 89%, respectively.

Finally, given the extensive degree of size heterogeneity at the apo(a) locus, it would have been expected that new apo(a) alleles would be encountered if a sufficient number of meioses were analyzed. In this sample, a total of 376 meiosis were examined and a single apo(a) allele was found in an offspring that was not present in either parent (Fig. 6). The fourth child, individual *f*, has *apo(a)16* and *apo(a)9*. Clearly, he inherited *apo(a)16* from his mother, but his father does not have *apo(a)9*. Paternity testing was performed using 7 unlinked varying number of tandem repeat (VNTRs), and in each case, the genotype of individual *f* was consistent with individual f being the child of individual a (39). The calculated probability of individual a not being the true father was $< 1 \times 10^{-6}$ (data not shown). Therefore, a mutation must have occurred in a paternal gamete which resulted in the generation of an apo(a) allele of different size.

Discussion

In this article we have evaluated the segregation of the apo(a) gene and plasma Lp(a) levels in 48 Caucasian families and found that virtually all the inter-individual variation in plasma Lp(a) levels was attributable to the genomic region encoding the apo(a) glycoprotein. It had been clear from previous family studies that plasma Lp(a) levels are largely genetically determined; prior estimates of the heritability of plasma Lp(a) levels have ranged from 0.75 to 0.98 (15, 17, 40, 41) which is comparable to our estimate of 0.85 (±8%). Initially, Lp(a) could only be detected in the plasma of Lp(a) of a third of individuals, and yet when family studies were performed, the inheritance pattern suggested a single autosomal dominant gene (11–13, 42–44). When more sensitive radioimmunoas-

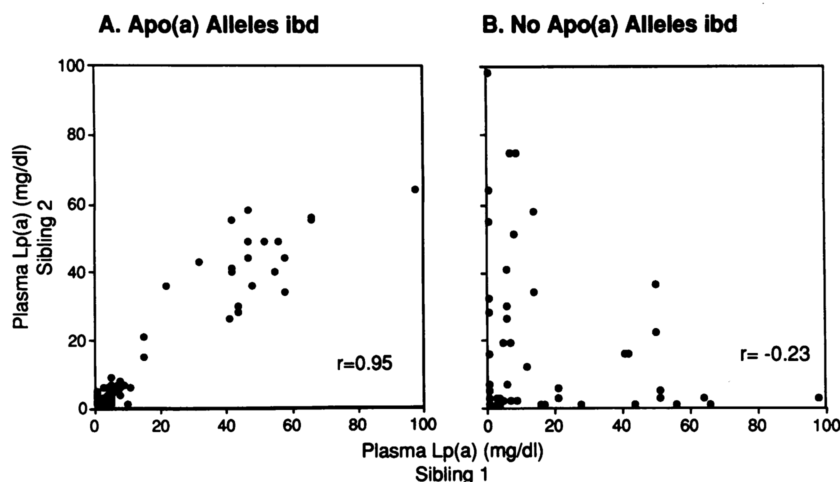

Figure 4. Scatter plot of Lp(a) levels for sibling pairs sharing (*A*) both ($n = 72$) or (*B*) no ($n = 52$) apo(a) alleles identical by descent.

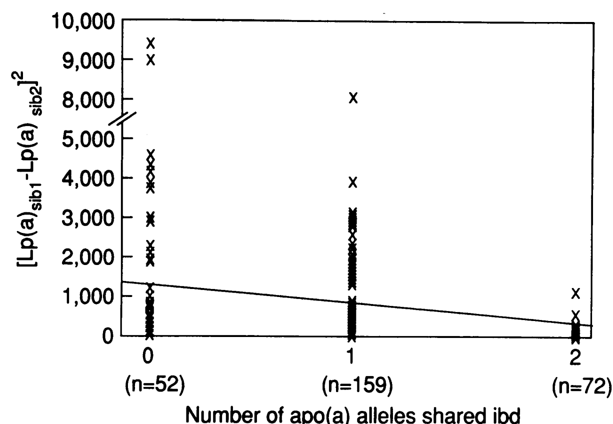

Figure 5. Squared difference between Lp(a) levels of siblings as a function of the proportion of apo(a) alleles shared identical by descent. The regression line for the squared difference on the proportion of apo(a) alleles shared identical by descent is given.

says were employed to measure plasma Lp(a) concentrations in families, there was evidence for a major gene, as well as polygenic factors, contributing to the plasma Lp(a) level (16, 17). In one large Caucasian pedigree, a major gene with three alleles was estimated to account for 73% of the variance in Lp(a) levels (16).

The first molecular clue that the apo(a) gene played a key role in the genetics of plasma Lp(a) concentrations, was the observation that the size of the apo(a) glycoprotein was inversely related to the plasma level of Lp(a) (19). Utermann and his colleagues estimated that differences in the size of the apo(a) glycoprotein accounted for 41% of the variation in inter-individual plasma Lp(a) levels (22). Further support for the apo(a) gene being the major gene influencing Lp(a) levels came from linkage analyses between segregation of plasma

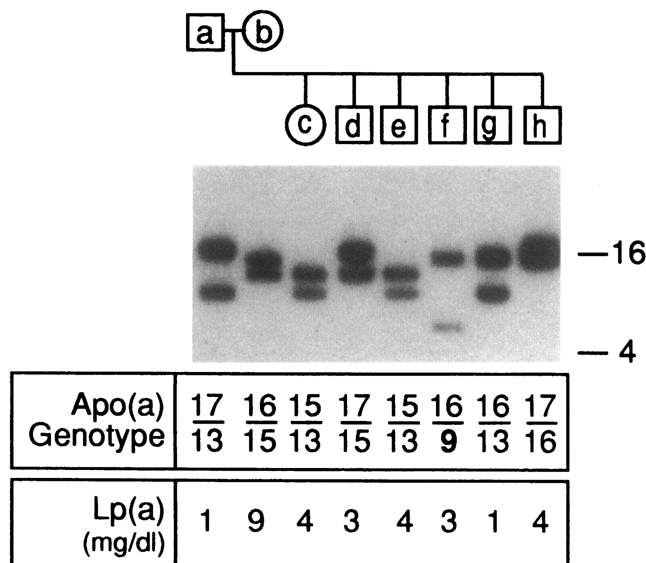

Figure 6. Genomic blot of kringle 4–encoding region of an apo(a) gene in a family in which there is generation of a new apo(a) allele of different length. Southern blotting of the kringle 4-encoding region of the apo(a) gene was performed as described in the Methods and Fig. 2. Individual *f* has inherited *apo(a)16* from his mother, and *apo(a)9* from his father. Paternity was confirmed by analysis of seven different VNTR sequences, as described in Methods.

Lp(a) concentrations and DNA sequences in the plasminogen gene which is closely linked to the apo(a) gene (45, 46).

The present study is distinguished from prior family studies by the fact that the apo(a) gene, rather than the expressed protein, was examined in relation to the level of Lp(a) in plasma. In prior studies the immunoblotting techniques used to examine apo(a) isoforms failed to detect protein products from all apo(a) alleles. Numerous exceptions to the inverse relationship between plasma Lp(a) levels and the size of the apo(a) protein were reported. It was suggested that these exceptions were due to the fact that not all apo(a) alleles were detected using the immunoblotting assay. In this study, apo(a) alleles associated with little or no production of apo(a) protein were included in the analysis. As a result, length variation within the kringle 4-encoding region of the apo(a) gene accounted for a greater proportion of the inter-individual variation in plasma Lp(a) concentrations than had been previously estimated (69% vs. 41% [22]).

The reason for the inverse correlation between the size of the apo(a) gene and the level of plasma Lp(a) is not known. Studies in primates have shown that there is not always a consistent relationship between the abundance of apo(a) mRNA, and its size, suggesting that differences in apo(a) gene transcription can not entirely account for this inverse relationship (26, 47). The size of the apo(a) mRNA transcript or glycoprotein may influence its rate of translation, or intracellular transport, respectively.

Alternatively, the observed inverse relationship may not be directly related to the number of kringle 4 repeats in the apo(a) gene, mRNA, or glycoprotein. The apo(a) alleles of different size might be in linkage disequilibrium with the actual sequences at the apo(a) locus that mediate the effect on plasma Lp(a) concentrations. The number of kringle 4 repeats in the apo(a) gene may not have a direct effect on plasma Lp(a) concentrations. In the marmoset monkey, for example, the plasma Lp(a) concentrations vary over a 100-fold range and yet there is only a single apo(a) isoform (48). In the current study, the contribution of the apo(a) gene was partitioned into two components to demonstrate that sequences at the apo(a) locus, other than the number of kringle 4 repeats, contribute importantly to plasma levels of Lp(a). If, however, the number of kringle 4 repeats in the apo(a) gene are in linkage disequilibrium with the actual sequences responsible for mediating the effect of the apo(a) gene or plasma level of Lp(a), then the contribution of the length polymorphism to the inter-individual variation in plasma Lp(a) levels has been overestimated.

Another possible cause for individuals with the same apo(a) genotypes having different plasma concentrations of Lp(a) is that alleles of the same size may differ in the composition of kringle 4 sequences. Not all the kringle sequences are identical. The first, as well as the last eight kringle repeats, differ from the common kringle 4 repeat (the so-called "A" repeat) by between 4 and 35 amino acids (25). Due to the frequent recombinational events involving this locus, is it highly likely that apo(a) alleles vary in their kringle 4 composition, as well as number. Subtle differences in the kringle 4 sequences may impact importantly on the synthesis, and/or degradation of Lp(a).

The length polymorphism in the apo(a) gene has a heterozygosity index comparable to that of number of tandem repeat (VNTR) loci employed in forensic and genetic linkage studies (39, 49, 50). The mutation rate at VNTR loci is several magni-

tudes higher than the usual bi-allelic DNA sequence polymorphisms (39, 51). Given the large number of different sized alleles at the apo(a) locus, a relatively high mutation rate was expected. Therefore, it was anticipated that mutations in the gene would be identified if a sufficient number of families were analyzed. We have observed one mutation of an apo(a) allele out of a total of 376 meioses, and this rate is of the same order of magnitude as the frequency of newly generated alleles for VNTR sequences (39, 51).

Most length polymorphisms in the human genome involve noncoding sequences. The coding regions of several mammalian genes have short tandem repeats (i.e., less the 50 basepairs) which are polymorphic in length (52–56). There are also examples of entire genes being tandemly repeated, as is the case with rDNA, 5S DNA, and the histone genes. The apo(a) length polymorphism is distinguished by the fact that the repeated sequence is large (5.5 kb) and contains both coding and noncoding sequences. The polyubiquitin gene (UbC), contains a large length polymorphism within its coding sequence, but all of the repeated sequences are contained in one exon, and each of the seven to nine repeats encodes the entire protein (57). The heterozygosity index of this length polymorphism is low (22%) compared to the apo(a) gene (94%). The extremely high degree of heterozygosity at the apo(a) locus may reflect the fact that it is under less selective pressure. A physiological function for this enigmatic protein has yet to be identified (2). Alternatively, there may be something intrinsic to the kringle 4-encoding sequences which make them more susceptible to recombinational events.

Mutations of repeated sequence domains result from either intrachromosomal or interchromosomal events. Initially, it was proposed that the mechanism primarily responsible for the high degree of size polymorphism in VNTRs was due to homologous recombination and unequal exchange during meiosis. However, molecular analysis of several new mutations revealed no exchange of flanking genetic markers, which suggests that intrachromosomal, rather than interchromosomal, events appear to be predominantly responsible (51, 58). Similarly, recent molecular analysis of tandem duplication within the Duchenne muscular dystrophy gene demonstrated that the recombinational events were due to intrachromosomal unequal exchange between sister chromatids rather than involving homologous chromosomes (59, 60). Efforts are now being directed to identify polymorphisms flanking the apo(a) gene to analyze the nature of the mutational event(s) responsible for the observed size heterogeneity at the apo(a) locus.

Acknowledgments

The authors thank Drs. Michael Brown and Joseph Goldstein for helpful discussions, Dr. Scott Grundy for assistance with the lipid and lipoprotein measures, and the laboratory of Dr. Steve Daiger at the University of Texas Health Science Center at Houston for markers which aided in the paternity analysis. Tommy Hyatt, Kathy Schueler, and Myriam Fornage provided excellent technical assistance.

This work was supported by grants from the Perot Family Foundation and grants, HL-47619, HL-20948, and HL-40613 from the National Institutes of Health, and 90-IJ-CS-0038 from the National Institute of Justice. E. Boerwinkle is a 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. Giulia Chiesa has a fellowship from the Italian Ministry of Scientific and Technological Research, and Carolin Lackner is a Schroedinger Scholar.

References

1. Berg, K. 1963. A new serum type system in man—the Lp system. *Acta Pathol. Microbiol. Scand.* 59:369–382.

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

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

4. 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.

5. Dahlen, G. H., J. R. Guyton, M. Attar, J. A. Farmer, J. A. Kautz, and A. M. Gotto, Jr. 1986. Association of levels of lipoprotein Lp(a), plasma lipids, and other lipoproteins with coronary artery disease documented by angiography. *Circulation.* 74:758–765.

6. Seed, M., F. Hoppichler, D. Reaveley, S. McCarthy, G. R. Thompson, E. Boerwinkle, and G. Utermann. 1990. Relation of serum lipoprotein(a) concentration and apolipoprotein(a) phenotype to coronary heart disease in patients with familial hypercholesterolemia. *N. Engl. J. Med.* 322:1494–1499.

7. Zenker, G., P. Költringer, G. Bonè, K. Niederkorn, K. Pfeiffer, and G. Jürgens. 1986. Lipoprotein(a) as a strong indicator for cerebrovascular disease. *Stroke.* 17:942–945.

8. Miles, L. A., G. M. Fless, E. G. Levin, A. M. Scanu, and E. F. Plow. 1989. A potential basis for the thrombotic risks associated with lipoprotein(a). *Nature (Lond.).* 339:301–303.

9. Edelberg, J. M., and S. V. Pizzo. 1991. Lipoprotein(a): the link between impaired fibrinolysis and atherosclerosis. *Fibrinolysis.* 5:135–143.

10. 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.

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

12. Sing, C. F., J. S. Schultz, and D. C. Shreffler. 1974. The genetics of the Lp antigen. II. A family study and proposed models of genetic control. *Ann. Hum. Genet.* 38:47–56.

13. Iselius, L., G. H. Dahlèn, U. De Faire, and T. Lundman. 1981. Complex segregation analysis of the Lp(a)/pre-β1-lipoprotein trait. *Clin. Genet.* 20:147–151.

14. Albers, J. J., and W. R. Hazzard. 1974. Immunochemical quantification of human plasma Lp(a) lipoprotein. *Lipids.* 9:15–26.

15. Hasstedt, S. J., D. E. Wilson, C. Q. Edwards, W. N. Cannon, D. Carmelli, and R. R. Williams. 1983. The genetics of quantitative plasma Lp(a): analysis of a large pedigree. *Am. J. Med. Genet.* 16:179–188.

16. Morton, N. E., K. Berg, G. H. Dahlèn, R. E. Ferrell, and G. Rhoads. 1985. Genetics of the Lp Lipoprotein in Japanese-Americans. *Genet. Epidemiol.* 2:113–121.

17. Hasstedt, S. J., and R. R. Williams. 1986. Three alleles for quantitative Lp(a). *Genet. Epidemiol.* 3:53–55.

18. Fless, G. M., M. E. Zum Mallen, and A. M. Scanu. 1986. Physiological properties of apolipoprotein(a) and lipoprotein(a-) derived from the dissociation of human plasma lipoprotein(a). *J. Biol. Chem.* 261:8712–8718.

19. 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.

20. Utermann, G., H. G. Kraft, H. J. Menzel, T. Hopferwieser, and C. Seitz. 1988. Genetics of the quantitative Lp(a) lipoprotein trait. I. Relation of Lp(a) glycoprotein phenotypes to Lp(a) lipoprotein concentrations in plasma. *Hum. Genet.* 78:41–46.

21. Utermann, G., C. Duba, and H. J. Menzel. 1988. Genetics of the quantitative Lp(a) lipoprotein trait. II. Inheritance of Lp(a) glycoprotein phenotypes. *Hum. Genet.* 78:47–50.

22. Boerwinkle, E., H. G. Menzel, H. G. Kraft, and G. Utermann. 1989. Genetics of the quantitative Lp(a) lipoprotein trait. III. Contribution of Lp(a) glycoprotein phenotypes to normal lipid variation. *Hum. Genet.* 82:73–78.

23. Sandholzer, C., D. M. Hallman, N. Saha, G. Sigurdsson, C. Lackner, A. Csàszàr, 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. (Abstr.)

24. Gaubatz, J. W., K. I. Ghanem, J. Guevara, Jr., M. L. Nava, W. Patsch, and J. D. Morrisett. 1990. Polymorphic forms of human apolipoprotein(a): Inheritance and relationship of their molecular weights to plasma levels of lipoprotein(a). *J. Lipid Res.* 31:603–613.

25. 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.

26. Hixson, J. E., M. L. Britten, G. S. Manis, and D. L. Rainwater. 1989.

Apolipoprotein(a) (apo(a)) glycoprotein isoforms result from size differences in apo(a) mRNA in baboons. *J. Biol. Chem.* 264:6013–6016.

27. Gavish, D., N. Azrolan, and J. L. Breslow. 1989. Plasma Lp(a) concentration is inversely correlated with the ratio of kringle IV/kringle V encoding domains in the apo(a) gene. *J. Clin. Invest.* 84:2021–2027.

28. Koschinsky, M. L., U. Beisiegel, D. Henne-Bruns, D. L. Eaton, and R. M. Lawn. 1990. Apolipoprotein(a) size heterogeneity is related to variable number of repeat sequences in its mRNA. *Biochemistry* 29:640–644.

29. 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.

30. Boerwinkle, E., and C. F. Sing. 1987. The use of measured genotype information in the analysis of quantitative phenotypes in man. III. Simultaneous estimation of the frequency and effects of the apolipoprotein E polymorphism and residual polygenic effects on cholesterol, betalipoprotein and triglyceride levels. *Ann. Hum. Genet.* 51:211–226.

31. Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA.* 81:1991–1995.

32. 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.

33. Lipid Research Clinic Program. 1982. Lipid and Lipoprotein Analysis: Manual of Laboratory Operations. Department of Health, Education and Welfare Publ NIH/75-628 Government Printing Office, Washington, DC.

34. Hasstedt, S. J., and P. Cartwright. 1981. PAP: Pedigree Analysis Program. Technical Report 13, Department of Medical Biophysics and Computing, University of Utah, Salt Lake City, UT.

35. Haseman, J. K., and R. C. Elston. 1972. The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.* 2:2–19.

36. Amos, C. I., R. C. Elston, A. F. Wilson, and J. E. Bailey-Wilson. 1989. A more powerful robust sib-pair test of linkage for quantitative traits. *Genet. Epidemiol.* 6:435–449.

37. Scheffe, H. 1959. The Analysis of Variance. John Wiley & Sons, Inc., New York. 221–235.

38. Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied Statistical Models. R. D. Irwin, Inc., Homewood, IL. 660–661.

39. Jeffreys, A. J., N. J. Royle, V. Wilson, and Z. Wong. 1988. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature (Lond.).* 332:278–281.

40. Albers, J. J., P. Wahl, and W. R. Hazzard. 1974. Quantitative genetic studies of the human plasma Lp(a) lipoprotein. *Biochem. Genet.* 11:475–486.

41. Hewitt, D., J. Milner, C. Breckenridge, and G. Macguire. 1977. Heritability of "sinking" pre-beta lipoprotein level: a twin study. *Clin. Genet.* 11:224–226.

42. Postle, A. D., J. M. Darmady, and D. C. Siggers. 1978. Double pre-β lipoprotein in ischaemic heart disease. *Clin. Genet.* 13:233–236.

43. Schultz, J. S., D. C. Shreffler, and C. F. Sing. 1974. The genetics of the Lp antigen. I. Its quantitation and distribution in a sample population. *Ann. Hum. Genet.* 38:39–46.

44. Hewitt, D., J. Milner, A. R. G. Owen, W. C. Breckenridge, G. F. Macguire, G. J. L. Jones, and J. A. Little. 1982. The inheritance of sinking-pre-beta lipoprotein and its relation to the Lp(a) antigen. *Clin. Genet.* 21:301–308.

45. Drayna, D. T., R. A. Hegele, P. E. Hass, M. Emi, L. L. Wu, D. L. Eaton, R. M. Lawn, R. R. Williams, R. L. White, and J.-M. Lalouel. 1988. Genetic linkage between lipoprotein(a) phenotype and a DNA polymorphism in the plasminogen gene. *Genomics.* 3:230–236.

46. Weitkamp, L. R., S. A. Guttormsen, and J. S. Schultz. 1988. Linkage between the loci for the Lp(a) lipoprotein (LP) and plasminogen (PLG). *Hum. Genet.* 79:80–82.

47. Azrolan, N., D. Gavish, and J. L. Breslow. 1991. Plasma lipoprotein(a) concentration is controlled by apolipoprotein(a) (Apo(a)) protein size and the abundance of hepatic apo(a) mRNA in a cynomolgus monkey model. *J. Biol. Chem.* 266(21):13866–13872.

48. Guo, H.-C., J.-B. Michel, Y. Blouquit, and M. J. Chapman. 1991. Lipoprotein(a) and apolipoprotein(a) in a new world monkey, the common marmoset (callithrix jacchus): Association of variable plasma lipoprotein(a) levels with a single apolipoprotein(a) isoform. *Arterioscler. Thromb.* 11:1030–1041.

49. Weber, J. L. 1990. Informativeness of human $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphisms. *Genomics.* 7:524–530.

50. Boerwinkle, E., W. Xiong, E. Fourest, and L. Chan. 1989. Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apolipoprotein B 3′ hypervariable region. *Proc. Natl. Acad. Sci. USA.* 86:212–216.

51. Jeffreys, A. J., R. Neumann, and V. Wilson. 1990. Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell.* 60:473–485.

52. Oberlè, I., F. Rousseau, D. Heitz, C. Kretz, D. Devys, A. Hanauer, J. Bouè, M. F. Bertheas, and J. L. Mandel. 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science (Wash. DC).* 252:1097–1102.

53. Azen, E., K. M. Lyons, T. McGonigal, N. L. Barrett, L. S. Clements, N. Maeda, E. F. Vanin, D. M. Carlson, and O. Smithies. 1984. Clones from the human gene complex coding for salivary proline-rich proteins. *Proc. Natl. Acad. Sci. USA.* 81:5561–5565.

54. Swallow, D. M., S. Gendler, B. Griffiths, G. Corney, J. Taylor-Papadimitriou, and M. E. Bramwell. 1987. The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. *Nature (Lond.).* 328:82–84.

55. Boerwinkle, E. S.-H. Chen, S. Visvikis, C. L. Hanis, G. Siest, and L. Chan. 1991. Signal peptide-length variation in human apolipoprotein B gene. *Diabetes.* 40:1539–1544.

56. McPhaul, M. J., M. Marcelli, W. D. Tiley, J. E. Griffin, R. F. Isidro-Gutierrez, and J. D. Wilson. 1991. Molecular basis of androgen resistance in a family with a qualitative abnormality of the androgen receptor and responsive to high-dose androgen therapy. *J. Clin. Invest.* 87:1413–1421.

57. Baker, R. T., and P. G. Board. 1989. Unequal crossover generates variation in ubiquitin coding unit number at the human UbC polyubiquitin locus. *Am. J. Hum. Genet.* 44:534–542.

58. Wolff, R. K., Y. Nakamura, and R. White. 1988. Molecular characterization of a spontaneously generated new allele at a VNTR locus: No exchange of flanking DNA sequence. *Genomics.* 3:347–351.

59. Hu, X., A. H. M. Burghes, D. E. Bulman, P. N. Ray, and R. G. Worton. 1989. Evidence for mutation by unequal sister chromatid exchange in the Duchenne muscular dystrophy gene. *Am. J. Hum. Genet.* 44:855–863.

60. Hu, X., P. N. Ray, and R. G. Worton. 1991. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:2471–2477.