Sequence Polymorphisms in the Apolipoprotein(a) Gene

Evidence for Dissociation between Apolipoprotein(a) Size and Plasma Lipoprotein(a) Levels

Jonathan C. Cohen, Giulia Chiesa, and Helen H. Hobbs
Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9046

Abstract

Apolipoprotein(a) [apo(a)], an apolipoprotein unique to lipoprotein(a) [Lp(a)], is highly polymorphic in size. Previous studies have indicated that the size of the apo(a) gene tends to be inversely correlated with the plasma level of Lp(a). However, several exceptions to this general trend have been identified. Individuals with apo(a) alleles of identical size do not always have similar plasma concentrations of Lp(a). To determine if these differences in plasma Lp(a) concentrations were due to sequence variations in the apo(a) gene, we examined the sequences of apo(a) alleles in 23 individuals homozygous for same-sized apo(a) alleles. We identified four single-strand DNA conformation polymorphisms (SSCPs) in the apo(a) gene. Of the 23 homozygotes, 21 (91%) were heterozygous for at least one of the SSCPs. Analysis of a family in which a parent was homozygous for the same-sized apo(a) allele revealed that each allele, though identical size, segregated with different plasma concentrations of Lp(a). These studies indicate that the apo(a) gene is even more polymorphic in sequence than was previously appreciated, and that sequence variations at the apo(a) locus, other than the number of kringle 4 repeats, contribute to the plasma concentration of Lp(a). (*J. Clin. Invest.* 1993. 91:1630–1636.) Key words: apolipoprotein(a) • atherosclerosis • DNA polymorphism • lipoprotein(a) • lipoproteins

Introduction

Lipoprotein(a) [Lp(a)][1] is distinguished from other plasma lipoproteins by the presence on its surface of a large glycoprotein, apolipoprotein(a) [apo(a)] (1, 2). Plasma levels of Lp(a) vary over a 1,000-fold range among individuals, but remain relatively constant over time in any given individual (3, 4). Family studies indicate that plasma levels of Lp(a) are largely genetically determined (5, 6). Boerwinkle et al. calculated that 90% of the interindividual variation in plasma Lp(a) levels in Caucasians is attributable to sequences at, or linked to, the apo(a) gene (6). The specific DNA sequences that determine the plasma Lp(a) concentration are currently being investigated.

One factor that contributes importantly to the plasma Lp(a) concentration is the size of the apo(a) glycoprotein (2, 7). The gene for apo(a) is highly polymorphic in size due to variation in the number of copies of an ~ 342-basepair (bp) tandem repeat. Each repeat encodes an ~ 114 amino acid cysteine-rich sequence that resembles the kringle 4 domain of plasminogen (8). The plasma levels of Lp(a) tend to be inversely related to the number of kringle 4 repeats in the apo(a) glycoprotein (2, 7).

Lackner et al. (9) developed a method to analyze the apo(a) gene structure directly by pulsed-field gel electrophoresis and genomic blotting. A total of 19 different apo(a) alleles (*apo(a)1–apo(a)19*) containing different numbers of kringle 4–encoding repeats could be distinguished. Among 103 unrelated Caucasians studied, 95 had two apo(a) alleles of different size, giving an overall heterozygosity index of 0.94. When this length polymorphism was used to analyze the co-segregation of apo(a) alleles and plasma Lp(a) levels in 40 nuclear families, it was found that siblings who inherited identical apo(a) alleles from their parents, i.e., apo(a) alleles identical by descent, had very similar plasma Lp(a) levels (6). In contrast, unrelated individuals who had apo(a) alleles of the same size sometimes had very different plasma concentrations of Lp(a). The observation that apo(a) alleles of the same size segregated with different levels of Lp(a) in different families suggested that apo(a) alleles of identical size may contain sequence variations that were responsible for the observed differences in plasma Lp(a) levels (6).

In a particularly informative family, the mother was homozygous for the same-sized apo(a) allele. However, analysis of the plasma Lp(a) levels in her offspring was not consistent with inheritance of identical maternal alleles. The present study was undertaken to identify sequence polymorphisms at the apo(a) locus that could be used to distinguish the two maternal apo(a) alleles in this family, as well as in other individuals with apo(a) alleles of identical size. To this end we used the PCR (10) and the single-strand DNA conformation polymorphism (SSCP) technique (11) and identified four different sequence polymorphisms in the apo(a) gene. In the family of interest, we showed that the two maternal alleles, though of identical size, segregate with very different plasma concentrations of Lp(a). Examination of the apo(a) alleles of 23 unrelated individuals whose apo(a) alleles were indistinguishable by pulsed-field gel analysis revealed that > 90% of their alleles differed in sequence in at least one of the polymorphic sites.

Methods

Pulsed-field gel electrophoresis. In 670 Americans (446 Caucasians [207 unrelated], 123 unrelated African Americans, and 101 unrelated Mainland Chinese individuals residing in America), apo(a) alleles were examined using a modification of the pulsed-field gel electrophoresis procedure described by Lackner et al. (9). Leukocytes isolated from whole blood using LeucoPrep tubes (Becton, Dickinson & Co., Lin-

---

Address reprint requests to Dr. Helen H. Hobbs, Department of Molecular Genetics, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-9046.
Received for publication 10 December 1992.

1. *Abbreviations used in this paper:* Lp(a), lipoprotein(a); SSCP, single-strand DNA conformation polymorphism.

J. Clin. Invest.
© The American Society for Clinical Investigation, Inc.
0021-9738/93/04/1630/07 $2.00
Volume 91, April 1993, 1630–1636

coln Park, NJ) were embedded in agarose plugs and stored at 4°C in 0.5 M Na_2EDTA, pH 8.0. The agarose-cellular plugs were washed in 1 × TE (1 mM Na_2EDTA, 10 mM Tris-Cl, pH 8.0) and then incubated twice for 2 h with 100 U KpnI (New England Biolabs, Beverly, MA) in the buffer suggested by the manufacturer. The agarose-cellular plugs were then subjected to transverse alternating gel electrophoresis for 18 h in a Gene Line II apparatus (Beckman Instruments Inc., Fullerton, CA). Lambda concatamers (Beckman Instruments Inc., Palo Alto, CA) were used as size standards. The size-fractionated DNA was transferred from the gel to a nylon membrane (ICN Biomedicals, Irvine, CA), and hybridized with an apo(a)-specific probe from the kringle 4–encoding region of the apo(a) gene, MP-1, as previously described (9). The filters were washed for 1 h and exposed to film with an intensifying screen (Lightening Plus, Dupont Co., Wilmington, DE) for 16 h at −70°C.

A total of 34, rather than 19, apo(a) alleles of different size could be resolved using this method (C. Lackner and H. H. Hobbs, unpublished observation). Therefore, the original nomenclature reported by Lackner et al. (9) was revised. In the present manuscript, each allele was designated by the total number of kringle 4–encoding repeats contained within its sequence. In our sample, the number of kringle 4–encoding sequences ranged from 12 to 51 (*apo(a)K-12* to *apo(a)K-51*).

ELISA assay of plasma Lp(a) concentrations. The plasma concentrations of Lp(a) were measured at GeneScreen (Dallas, TX) using a sandwich ELISA assay exactly as previously described (9). In this assay, the first antibody is a rabbit anti-human Lp(a) antibody and the detection antibody was IgG1A[2], an antibody directed against the common kringle 4 repeat (provided by Dr. Gerd Utermann, Innsbruck, Austria).

Amplification and sequencing of specific apo(a) introns. Oligonucleotides complementary to unique sequences from the apo(a) cDNA (8) were synthesized using ABI 380B and 394 DNA synthesizers (Applied Biosystems, Inc., Foster City, CA) (Table I). To amplify apo(a) introns between specific kringle 4–encoding sequences, 25-base oligonucleotides were designed such that the 3′ sequence of the oligonucleotide differed from the sequence of the other apo(a) and plasminogen kringle 4–encoding sequences (12).

Pairs of oppositely oriented 25-base oligonucleotides (Table I) were used to amplify sequences from genomic DNA using PCR (10). Each PCR reaction was performed in a 20-μl volume containing 0.1 μg genomic DNA, 20 pmol of each primer, 2 nmol dNTP, and 2 U of *Thermus aquaticus* DNA polymerase (Promega Corp., Madison, WI) in the buffer supplied by the manufacturer. The PCR mixtures were overlaid with mineral oil and subjected to 30 cycles of denaturation, annealing, and extension in a thermocycler (Perkin-Elmer Cetus, Norwalk, CT) using the conditions described in the legend to Fig. 1.

To confirm the specificity of the PCR amplifications, each PCR amplified fragment was sequenced as described previously (13, 14). An oligonucleotide with sequence identity to an apo(a) exon (Table I) was end-labeled with [γ-^{32}P]ATP (6,000 Ci/mmol, Amersham Corp., Arlington Heights, IL) and used as a sequencing primer. The sequencing reactions were performed in 10-μl volumes containing 1 pmol sequencing primer, 0.1 μg DNA template, 1 μl 10× PCR buffer (10× PCR buffer = 500 mM KCl, 15 mM $MgCl_2$, 100 mM Tris-Cl, pH 8.3), 5 μl termination mix, and 0.25 U *Thermus aquaticus* DNA polymerase. The termination mixes contained C mix (20 μM dCTP, dTTP, dATP, dGTP, and 400 μM ddCTP), T mix (20 μM dCTP, dTTP, dATP, dGTP, and 800 μM ddTTP), A mix (20 μM dCTP, dTTP, dATP, dGTP, and 600 μM ddATP), and G mix (20 μM dCTP, dTTP, dATP, dGTP, and 200 μM ddGTP). The reactions were overlaid with mineral oil and subjected to 30 cycles of amplification in a thermal cycler. The temperatures for the first 20 cycles were 95°C for 1 min, 55°C for 1 min, and 70°C for 1 min; for the last 10 cycles the 55°C annealing step was omitted. The four reactions were terminated by addition of 10 μl formamide dye (98% formamide [vol/vol], 0.1% [wt/vol] bromophenol blue, 0.1% [wt/vol] xylene cyanol), denatured by heating to 95°C for 3 min, and 3 μl was loaded onto an 8% denaturing polyacrylamide gel in 1 × TBE (1 × TBE = 0.05 M Tris, 0.05 M borate, 10 mM Na_2EDTA). The sequencing gel was run for 2 h at 75 W, dried, and then exposed to XAR-5 film (Eastman Kodak Co., Rochester, NY) for 3 h at −70°C.

Identification of SSCPs. The PCR-amplified introns were screened for SSCPs using a modification of the method of Orita et al. (11). The PCR-amplified DNA fragments were too large (> 1,400 bp) to be analyzed directly by SSCP. Therefore, for each fragment a restriction enzyme was identified that generated cleavage products between 200 and 400 bp. To perform the SSCP analysis, each fragment was amplified by PCR in a total volume of 20 μl containing 2 μl 10× PCR buffer, 1 nmol of dNTP, 20 pmol of each 25-bp flanking primer (Table I), 0.1 μg of genomic DNA, 3.3 pmol [α-^{32}P]dCTP (3,000 Ci/mmol) and 2 U of *Thermus aquaticus* DNA polymerase. The fragments were amplified in a thermocycler as described in the legend to Fig. 1. After amplification, 10 μl of the reaction was digested in a total volume of 50 μl at 37°C for 2 h with 5 U of the appropriate restriction enzyme in a buffer supplied by the manufacturer (New England Biolabs). 5 μl of the digestion reaction was diluted in 20 μl of formamide dye, and denatured at 95°C for 3 min. 3 μl of the denatured sample was loaded onto a 6% nondenaturing polyacrylamide gel in 2 × TBE with 10% glycerol and subjected to electrophoresis at 300 V for 14 h. The gel was dried and exposed to Kodak XAR-5 film for 12 h at −70°C.

Table I. Oligonucleotide Primers for Amplification of Specific Apo(a) DNA Sequences

Location	Oligo no.	Sequence
1. 5′ flanking region*	LP1-5′	5′-TGACATTGCACTCTCAAATATTTTA-3′
	LP-3′	CATATACAAGATTTTGAACTGGGAA
2. K4-1‡	GC1C	TTGGTCATCTATGACACCACATCAA
	3CA3′	GGCTCGGTTGATTTCATTTTTCAGC
3. K4-33/34§	33a	CTTGGTGTTATACCATGGATCCCAA
	34c	CACCATGGTAGCAGTCCTGGACTGT
4. K4-35/36‖	35e	ATCGAGTGTCCTCACAACTCCCACA
	36b	GGGGTCCTCTGATGCCAGTGTGGTA

* The polymorphism is located in the 5′ untranslated region. Base number 99 (using the numbering system of Ichinose [15]) is an A in 14% of alleles, and a G in 86% of alleles. In addition, an infrequent polymorphism can be seen by SSCP but the polymorphic base has not been identified. ‡ The polymorphism is located in the intron between the two exons that encode the first kringle 4 repeat (K4-1). § The polymorphism is located in the intron between the second exon of K4-33 and the first exon of K4-34. ‖ The polymorphism is located in the intron between the second exon of K4-35 and the first exon of K4-36.

Results

Identification of four SSCP polymorphisms at the apo(a) locus. PCR-based analysis of specific kringle 4 repeats in the apo(a) gene is complicated by the presence of multiple highly homologous kringle 4–encoding sequences in the apo(a) and plasminogen genes, as well as other members of the plasminogen gene family. In the only human apo(a) cDNA that has been sequenced to date, it was estimated that ~ 25 of the kringle 4 repeats were identical in sequence (the so-called K4-A repeats) (8). The first (K4-1) and the last eight of the 37 kringle 4–encoding sequences (K4-30 through K4-37) differed in sequence from the common repeat by 11–71 nucleotides. In order to develop oligonucleotides that specifically amplify only a single sequence in the apo(a) gene, the sequences of all the kringle 4 repeats were compared to each other as well as to the kringle 4–

encoding sequence of plasminogen (8). Oligonucleotide primers were synthesized so that the 3′ nucleotide sequence was specific for the kringle 4 sequence targeted for amplification (see Table I) (12). To ensure that sequences from the

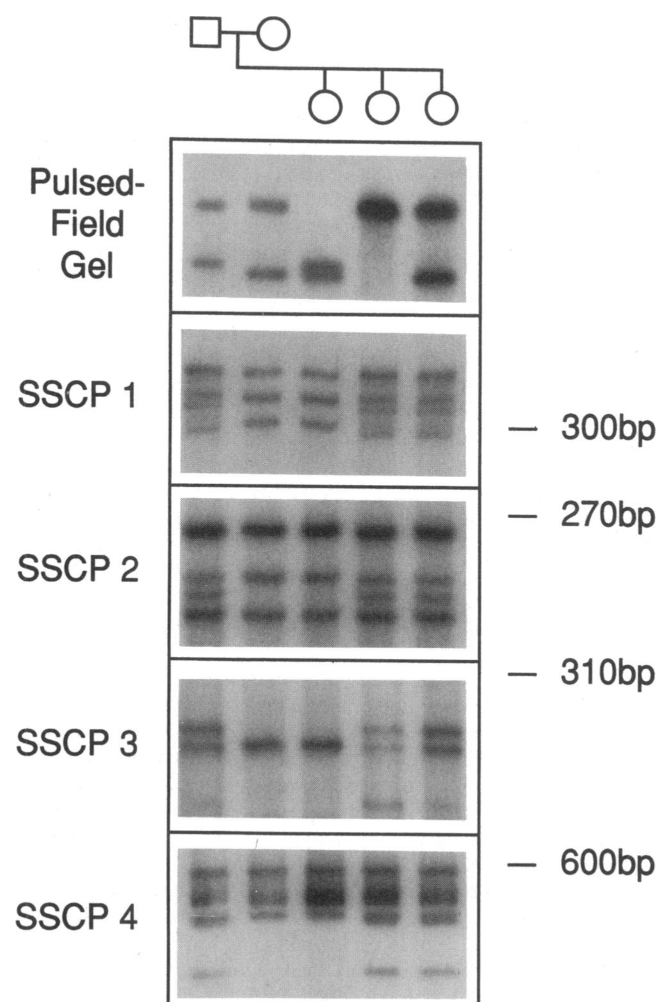

Figure 1. Segregation of four SSCPs at the human apo(a) locus in a nuclear family. For each polymorphism, 25-base oligonucleotides (Table I) and 0.1 μg of genomic DNA were used to amplify by PCR the intervening sequence from each family member (10). SSCP 1 was amplified using 30 cycles of 95°C for 1 min, 55°C for 1 min, and 72°C for 1 min. SSCP 2, 3, and 4 were amplified using 30 cycles of 95°C for 1 min and 68°C for 3 min. The amplification products were denatured and size-fractionated on 6% nondenaturing polyacrylamide gels (unless otherwise specified). The gels were dried and then exposed to XAR-5 film (Eastman Kodak Co.) for 12 h at −70°C. SSCP 1: ~ 300-bp fragment from the 5′ flanking region was amplified by PCR using oligonucleotides LP1-5′ and LP1-3′. The alleles bearing polymorphic sites designated 1 (*upper band*) and 2 (*lower band*) segregate in this family. SSCP 2: the intron separating the exons encoding kringle 4-1 (K4-1) was amplified by PCR using oligonucleotides GC-1C and 3CA3′ and digested with HaeIII before gel electrophoresis as described in the Methods. SSCP3: the intron between the second exon of kringle 4-33 and the first exon of kringle 4-34 (K4-33/34) was amplified using oligonucleotides 33a and 34c, digested with HinfI, and size-fractionated on an 8% nondenaturing polyacrylamide gel. SSCP4: the intron between the second exon of kringle 4-35 and the first exon of kringle 4-36 (K4-35/36) was amplified with oligonucleotides 35e and 36b and digested with HinfI before gel electrophoresis.

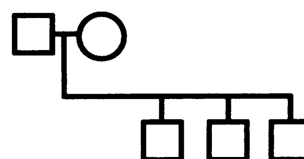

Figure 2. Segregation of a third SSCP band from the apo(a) 5′ flanking region in a nuclear family. The intervening sequence between 25-base oligonucleotides (Table I) was amplified by PCR from each family member as described in Fig. 1. The upper band in this figure is identical with the lower band (band *2*) of SSCP 1 shown in Fig. 1. The other band (band *3*) thus migrates faster than bands *1* and *2*.

specific kringle were the only ones amplified, the PCR product was sequenced. In each case, the sequence of the product confirmed specific amplification of the targeted kringle-4 repeat (data not shown).

With these techniques, four polymorphic sites (1–4) were identified using genomic DNA from six unrelated Caucasians. The four sites were located in the following regions: (*a*) the 5′ flanking region, (*b*) the intron separating the two exons that encode the first kringle 4 repeat (K4-1), (*c*) the intron separating kringle 33 from 34 (K4-33/34), and (*d*) the intron separating kringle 35 from 36 (K4-35/36).

Analysis of the segregation of the four sequence polymorphisms in three nuclear families demonstrated that they each were inherited in a co-dominant Mendelian fashion and co-segregated with the apo(a) alleles determined by pulsed-field gel electrophoresis. The analysis of one such family is shown in Fig. 1. The allele generating the faster and slower migrating fragments have been arbitrarily designated (2) and (1), respectively. The three intronic polymorphisms were biallelic.

The polymorphic site in the 5′ flanking region had three alleles. This polymorphic site was revealed using specific oligonucleotides that were complementary to the 5′ flanking sequence of the *apo(a)I* gene as reported by Ichinose (15). Though it has not been unequivocally established that the *apo(a)I* gene is the functional apo(a) gene, a sequence reported by Ichinose (15) differs by only a single bp from the 5′ untranslated sequence reported by McLean et al. (8). Base number 99 is a G in the sequence reported by Ichinose (15) and an A in the sequence published by McLean et al. (8). To determine that the sequence we have amplified is in the 5′ flanking region of the apo(a) gene, PCR-amplified polymorphic fragments were each sequenced and band 1 had an A and band 2 a G at position 99 (data not shown). A third SSCP band was identified and is shown in Fig. 2; the sequence difference responsible for this third band was not identified.

The heterozygosity indices for the SSCPs were determined by analyzing 25 unrelated Caucasians and were as follows: (*a*) 5′ flanking region, 24%; (*b*) kringle 4-1, 46%; (*c*) kringle 4-33/34, 44%; and (*d*) kringle 4-35/36, 42%. All of the polymorphisms were found to be in Hardy-Weinberg equilibrium. 23 (92%) of these 25 individuals were heterozygous for at least one of the four polymorphisms. Each of these polymorphisms was

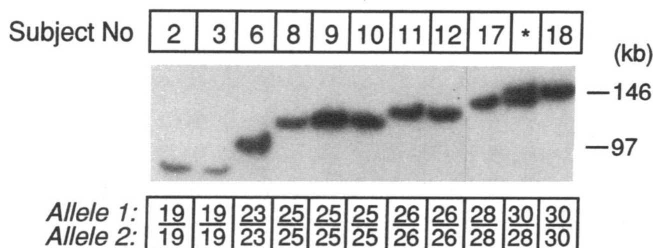

Figure 3. Pulsed-field gel electrophoresis of 10 individuals with apo(a) alleles of the same size. High molecular weight genomic DNA was digested with KpnI and the resultant fragments were size-fractionated using pulsed-field gel electrophoresis, transferred to a nylon membrane, and hybridized with an apo(a) specific probe (MP-1) as described in Methods. One of the 11 individuals examined (who is indicated by an asterix) was heterozygous for *apo(a)K-28* and *apo(a)K-30.*

also identified in African-American and Chinese individuals, although the heterozygosity indices were not determined in these ethnic groups.

Identification of SSCP polymorphisms in individuals with apo(a) of identical size. The kringle 4–encoding region of the apo(a) gene was analyzed in 670 individuals using pulsed-field gel electrophoresis and Southern blotting (9). There were 49 individuals (27 Caucasians [6% of the number analyzed], 12 African-Americans [10%], and 10 Chinese [10%]) whose apo(a) alleles could not be distinguished by size-fractionation (data not shown). Pulsed-field gel analyses of a representative sample of these are shown in Fig. 3. 23 individuals homozygous for the same-sized apo(a) alleles were randomly selected and screened for the presence of the four SSCPs and the results of this analysis are shown in Table II. Of the 23 individuals studied, a total of 21 were heterozygous for at least one of the four SSCPs. Therefore, the overall heterozygosity index was as high in the individuals who were homozygous for the same-sized apo(a) alleles as had been seen in the random sample.

For each polymorphic site, individuals with the same SSCP pattern had plasma Lp(a) levels that differed over a wide range. For example, Caucasian individuals homozygous for band 2 at SSCP 1 (i.e., individuals 3, 5–8, 11, and 14) had plasma Lp(a) levels that ranged from 3 mg/dl (individual 11) to 41 mg/dl (individual 5). African-American individuals who were homozygous for the same band (individuals 1, 4, and 9) had plasma Lp(a) concentrations that ranged from 29 to 90 mg/dl. For each of the other three polymorphisms, there was no association between the SSCP pattern and the plasma concentrations of Lp(a). Therefore, none of the sequence polymorphisms were associated with differences in the plasma concentrations of Lp(a).

Nor was there any association between the SSCP haplotypes and the plasma concentrations of Lp(a). In the subset of

Table II. SSCPs in Individuals with Two Apo(a) Alleles of Identical Size

No.	Ethnic background	Apo(a)* allele	Plasma Lp(a)	SSCP 1	SSCP 2	SSCP 3	SSCP 4
			mg/dl				
1	**AA**	**18**	**90**	**22**	**12**	**22**	**22**
2	CC	19	84	12	12	22	22
3	CC	19	7	22	12	22	11
4	AA	19	29	22	12	11	12
5	CC	21	41	22	12	12	12
6	CC	23	33	22	22	22	22
7	CC	23	51	22	11	12	12
8	**CC**	**25**	**7**	**22**	**12**	**22**	**22**
9	**AA**	**25**	**65**	**22**	**12**	**22**	**22**
10	AA	25	90	11	12	11	11
11	**CC**	**26**	**3**	**22**	**12**	**22**	**22**
12	CC	26	3	23	22	12	12
13	CC	27	< 1	12	na	22	22
14	**CC**	**28**	**16**	**22**	**12**	**22**	**22**
15	CC	28	10	23	22	12	12
16	CC	28	5	12	12	12	12
17	AA	28	21	11	22	11	11
18	CC	30	30	12	22	12	12
19	CH	33	5	12	22	22	22
20	CH	33	14	12	22	11	12
21	CC	33	1	11	12	22	22
22	CH	36	3	12	22	12	12
23	AA	37	4	12	12	22	22

Abbreviations: The numbers refer to numbers used in Table I; na, not ascertained; CC, Caucasian; AA, African-American; CH, Chinese. * The apo(a) alleles are designated by numbers based on the size of a KpnI restriction fragment from the kringle 4–encoding region of the apo(a) gene, as previously described (9).

23 individuals who were homozygous for the same-sized apo(a) gene, a total of eight different haplotypes (A–H) could be determined unequivocally by examining the four SSCPs (Table III). Since the families of the individuals were not studied, the haplotypes could only be determined in those individuals homozygous for at least three of the sequence polymorphisms. For example, individuals 1, 8, 9, 11, and 14 all have one apo(a) allele of SSCP haplotype A and one of SSCP haplotype B (shown in boldface type in Table II). Despite having apo(a) genes with identical haplotypes, they had plasma concentrations of Lp(a) that ranged from 3 to 16 mg/dl in the Caucasian subgroup and from 65 to 90 mg/dl in the African-Americans. Individuals 8 and 9 not only had identical SSCP haplotypes, but were also homozygous for the same-sized apo(a) alleles (*apo(a)K-25*). Their plasma Lp(a) levels were widely disparate (7 vs. 65 mg/dl). However, direct comparison between these two individuals may not be appropriate, since they belong to different ethnic groups (Caucasian and African-American). Previous studies have shown African-Americans have plasma concentrations of Lp(a) that are on average two- to threefold higher than those of Caucasians (16), and it has not been determined whether these differences in plasma levels of Lp(a) are due to sequence differences at the apo(a) locus, or to the effect of other genetic or environmental factors. Therefore, it is possible that the widely different plasma concentrations of Lp(a) in individuals 8 and 9 (Table II) are due to factors other than differences in the sequences of their apo(a) genes.

Many apo(a) alleles of the same size were associated with more than one SSCP haplotype (Table III). For example, three individuals were homozygous for *apo(a)K-25* (individuals 8–10, Table II). Amongst the six *apo(a)K-25* alleles, there were four different apo(a) SSCP haplotypes (A, B, C, and H) (Table III). These findings indicate that apo(a) alleles of the same size differ in sequence and that the number of different apo(a) alleles is much larger than estimated based on the size of the apo(a) gene. At least 34 apo(a) alleles can be resolved by size fractionation using pulsed-field gel electrophoresis (unpublished observations). If each of these alleles of different size actually comprises three or more different alleles, then there are over 100 apo(a) alleles in the human population.

Table III. SSCP Haplotypes in 11 Individuals Homozygous for the Same-sized Apo(a) Alleles in Whom the Apo(a) Haplotypes Are Known

Haplotype no.	SSCP haplotype*	No. alleles with haplotype	Apo(a) alleles‡ with haplotype
		n (%)	
A	2, 2, 2, 2	8 (33)	18, 23(2), 25(2), 26, 28, 33
B	2, 1, 2, 2	5 (21)	18, 25(2), 26, 28
C	1, 2, 1, 1	3 (12)	25, 28(2)
D	1, 2, 2, 2	2 (8)	28, 33
E	1, 1, 2, 2	1 (12)	33
F	2, 1, 2, 1	1 (4)	19
G	2, 2, 2, 1	1 (4)	19
H	1, 1, 1, 1	1 (4)	25

* SSCP haplotypes were constructed using SSCP 1, 2, 3, and 4, as described in Results. ‡ Apo(a) alleles (*apo(a)K-12-apo(a)K-51*) as determined by pulsed-field gel electrophoresis (9).

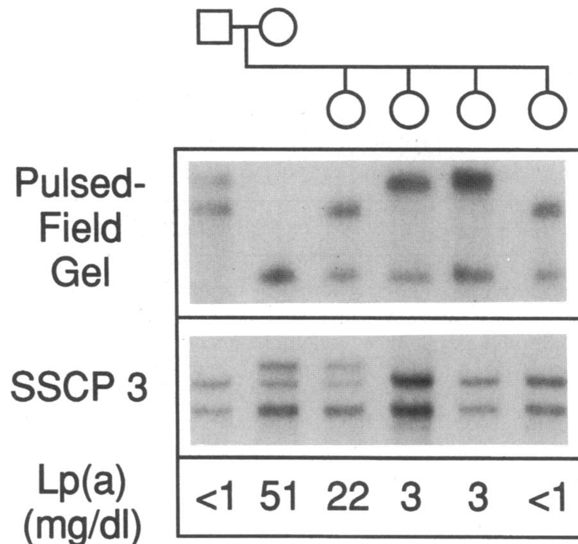

Figure 4. Segregation of plasma Lp(a) levels with apo(a) alleles in a nuclear family. Plasma Lp(a) levels were determined by sandwich ELISA as described previously (9). Apo (a) allele sizes were determined by pulsed-field gel electrophoresis and genomic blotting as described in Methods. The SSCP polymorphism was identified by PCR amplification of the intron between kringles 33 and 34 using oligonucleotides described in Table I. PCR conditions were 30 cycles of 95°C for 1 min and 68°C for 3 min. Samples were digested with HinfI, denatured, and subjected to electrophoresis on an 8% nondenaturing polyacrylamide gel.

Analysis of a family in which there is dissociation between the apo(a) genotype and the plasma concentration of Lp(a). We have previously identified a nuclear family in which the size polymorphism revealed by pulsed-field gel electrophoresis was not informative (Fig. 4) (9). The mother was homozygous for the *apo(a)23* allele and had a plasma Lp(a) concentration of 50 mg/dl. The father had *apo(a)K-30* and *apo(a)K-33* and a very low (< 1 mg/dl) level of plasma Lp(a), so neither of his alleles should have contributed appreciably to the plasma Lp(a) levels in the offspring. Each of their four daughters inherited a copy of *apo(a)K-23* from their mother. Three of the offspring had very low plasma Lp(a) levels (< 1–3 mg/dl), while the other child had a significantly higher plasma level of Lp(a) (22 mg/dl) (Fig. 4). The apo(a) alleles from the mother were examined for the presence of the four sequence polymorphisms. She was found to be heterozygous for the polymorphism in K4-33/34 (i.e., three in Tables I and II). By analyzing this polymorphism, the two maternal *apo(a)K-23* alleles could be distinguished and their segregation followed. The three daughters who had low plasma Lp(a) concentrations all inherited the same *apo(a)K-23* allele from their mother, whereas the daughter with the high plasma level of Lp(a) inherited the other *apo(a)K-23* allele. The two *apo(a)K-23* alleles of the mother thus were associated with very different plasma levels of Lp(a).

Discussion

The results of the present study demonstrate that each size-class of apo(a) alleles identified by pulsed-field gel electrophoresis comprises several distinct alleles that differ in sequence.

Analyses of the apo(a) genes in 23 individuals whose apo(a) alleles were indistinguishable using the pulsed-field gel electrophoresis technique revealed that 21 (91%) were heterozygous at one or more sites in the apo(a) gene. The sequence heterogeneity identified in the present study indicates that the apo(a) gene is even more polymorphic than was originally determined. Initially 19 (9), and now 34 apo(a) alleles can be distinguished by pulsed-field gel electrophoresis (C. Lackner and H. H. Hobbs, unpublished observation). For each apo(a) allele that has been analyzed by SSCP, there are at least two and usually multiple apo(a) alleles that can be distinguished as having different haplotypes by examination of the four SSCPs reported in this paper. Therefore, if the pulsed-field data is combined with the SSCP data, it can be estimated that there are more than 100 different alleles at the apo(a) locus.

These findings have important implications for the biology of Lp(a). The apo(a) glycoprotein and gene sizes tend to be inversely related to the plasma Lp(a) level, and a causal relationship between the size of the apo(a) gene and the plasma level of Lp(a) has been suggested (2, 7, 17). However, individuals with the same-sized apo(a) allele may have very different plasma levels of Lp(a) (6, 9). By using the SSCP polymorphisms to analyze a particularly informative family in which the mother was homozygous for the same-sized apo(a) allele and yet her offspring had very different plasma levels of Lp(a), each of her apo(a) alleles, though of identical size could be shown to c cosegregate with either a high (i.e., 22 mg/dl) or low ($\leq$ 3 mg/dl) plasma level of Lp(a). These findings indicate that sequence variations at (or linked to) the apo(a) gene that are unrelated to the number of kringle 4 repeats can cause differences in plasma Lp(a) concentrations. The reason why the same apo(a) allele in the mother is associated with a higher plasma concentration of Lp(a) than in the daughter has not been determined. The mother, aged 33 yr, is healthy and premenopausal. The daughter, aged 10 yr, is also healthy. We are in the process of locating additional maternal relatives so we can further analyze this allele.

The fact that apo(a) is so polymorphic in size and sequence has complicated the development of assays to measure plasma Lp(a) concentrations (18). The detecting antibody employed in the ELISA assay used in this study is directed against an epitope in the kringle 4 domain (19). It has not been conclusively demonstrated that this antibody reacts equally to all the different apo(a) isoforms (19). To address this problem, we have developed an antibody to a unique epitope within the protease domain of apo(a). When this antibody was used for immunoblotting, the relative immunoreactivity of the different apo(a) isoforms was similar to that seen using the kringle 4–specific antibody (data not shown). This suggests that our ELISA measurements of plasma Lp(a) levels reflect true differences in concentration and are not an artifact of the antibody employed.

In the small sample of unrelated individuals analyzed in this study there appeared to be no association between the polymorphic sequences and the plasma concentrations of Lp(a). This observation is somewhat surprising given the fact that most of the interindividual variation in plasma Lp(a) levels can be attributed to the apo(a) gene (6). A possible reason for the observed lack of association between SSCP patterns and plasma concentrations of Lp(a) is that extensive recombination has occurred within the apo(a) gene. To support this interpretation, each of the four sequence polymorphisms was identified in apo(a) alleles of different sizes. For example, the most common haplotype, haplotype A, was found in alleles containing 18, 23, 25, 26, 28, and 33 kringle 4 repeats (Table III). While the presence of identical sequence polymorphisms in different sized apo(a) alleles may be due to recurrent mutations at the specific sites associated with each SSCP, the high heterozygosity indices of these polymorphisms in the Caucasian population, and the fact that each of the polymorphisms is present in other population groups makes this explanation extremely unlikely. A more probable explanation is that the mutations causing these sequence polymorphisms occurred early in the lineage of *Homo sapiens* (before the divergence of the main racial groups), and that the sequence polymorphisms became associated with different sized apo(a) alleles by subsequent loss and/or gain of kringle 4 repeats from the primordial apo(a) allele in which they arose.

Variation in the copy number of tandemly repeated sequences is thought to arise by the misalignment and subsequent unequal exchange of repetitive DNA segments during recombination between chromosomes in meiosis or mitosis (20, 21). The polymorphisms described in this study offer an indication of the types of recombination that have occurred at the apo(a) locus. The SSCPs reported here flank the subset of identical kringle 4 repeats. It has been proposed that recombinational events that generate the size polymorphism in the apo(a) gene occur within this region (8). The polymorphisms in the 5′ flanking region of the apo(a) gene and in the intron between kringles 35 and 36 are found in all possible phases (1,1 ; 1,2 ; 2,1 ; and 2,2). This distribution of polymorphic markers could not have arisen by only intrachromosomal recombinational events. Therefore, at least one recombinational event between nonidentical chromosomes (i.e., interchromosomal rearrangement causing phase switching) has occurred between these sites. This finding does not mean that the high degree of size polymorphism at the apo(a) locus can be solely attributable to interchromosomal recombinational events.

The SSCPs described in the study can be employed to determine the segregation of apo(a) alleles in families in which pulsed-field gel electrophoresis is uninformative. We have analyzed five nuclear families in which the apo(a) alleles of one of the parents were indistinguishable by pulsed-field gel electrophoresis, and in four of these families the parental alleles could be distinguished using these SSCPs (data not shown). Furthermore, detailed analysis of these polymorphisms in larger groups of individuals may provide insights into the region(s) of the apo(a) gene that determine plasma Lp(a) levels.

Acknowledgments

We wish to acknowledge the excellent technical assistance of Lisa Krijger, Tommy Hyatt, Kathy Schueler, and Carla Leffert. We thank Eric Boerwinkle and Allan Gaw for helpful discussions.

This work was supported by National Institutes of Health grants (HL-20948, HL-47619), and the Perot Family Fund. Giulia Chiesa is supported by a grant from the Italian Ministry of Scientific and Technological Research. H. H. Hobbs is a principle investigator of the American Heart Association.

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

5. 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.

6. 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.

7. 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.

8. 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.

9. 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.

10. Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335–350.

11. Orita, M., Y. Suzuki, T. Sekiya, and K. Hayashi. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics.* 5:874–879.

12. Ugozzoli, L., and R. B. Wallace. 1991. Allele-specific polymerase chain reaction. In Methods: A Companion to *Methods Enzymol.* 2:42–48.

13. Cohen, J. C., J. J. Cali, D. F. Jelinek, M. Mehrabian, R. S. Sparkes, A. J. Lusis, D. W. Russell, and H. H. Hobbs. 1992. Cloning of the human cholesterol 7α-hydroxylase gene (CYP7) and localization to chromosome 8q11-q12. *Genomics.* 14:153–161.

14. Lee, J.-S. 1991. Alternative dideoxy sequencing of double-stranded DNA by cyclic reactions using Taq polymerase. *DNA Cell Biol.* 10:67–73.

15. Ichinose, A. 1992. Multiple members of the plasminogen-apolipoprotein(a) gene family associated with thrombosis. *Biochemistry.* 31:3113–3118.

16. Guyton, J. R., G. H. Dahlen, 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.

17. 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.

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

19. 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.

20. Smith, G. P. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science (Wash. DC)*. 191:528–535.

21. Wolff, R. K., R. Plaetke, A. J. Jeffreys, and R. White. 1989. Unequal crossing-over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. *Genomics.* 5:382–384.