Am J Hum Genet 32:339–347, 1980

Genetics of the Apolipoprotein E System in Man

GERD UTERMANN,[1] ULRICH LANGENBECK,[2] ULRIKE BEISIEGEL,[1] AND WILFRIED WEBER[1]

SUMMARY

The polymorphism of apolipoprotein E (Apo E) in man is controlled by two codominant alleles, *Apo* E^n and *Apo* E^d, at the *Apo E-N/D* locus and by two alleles, the dominant, *Apo* $E4^+$, and the recessive, *Apo* $E4^0$, at the *Apo E4* locus.

Frequency distribution analysis of Apo E phenotypes demonstrated a highly significant association between both systems ($P \sim 1\%$). The Apo E4-(+) variant was about twice as frequent in phenotype Apo E-N (30.1%) than in phenotype Apo E-ND (16.4%). The phenotypic combination Apo E-D/-E4(+) was not observed. The segregation of Apo E phenotypes in informative matings is consistent with a close linkage of both loci.

The results may be explained by different models. On the basis of the present data, these models cannot be distinguished by formal genetic criteria. (1) Haplotypes *Apo* $E^n/E4^+$, *Apo* $E^n/E4^0$, and *Apo* $E^d/E4^0$ determine the different phenotypes, and a linkage disequilibrium exists of $\Delta = .0147$ between the *E-N/D* and *E4* loci. (2) The fourth haplotype, *Apo* $E^d/E4^+$, exists, but the gene $E4^+$ is not expressed in coupling with *Apo* E^d. The four-haplotype model seems more attractive in view of Apo E-N/D polymorphism's quantitative character and of biochemical results, which show that phenotypes Apo E-N and Apo E-D differ in the apparent molecular weight (M_r) of the respective major Apo E polymorphic form. Hence, the *Apo E-N/D* locus may control structural genes involved in the posttranslational modification of Apo E. (3) Finally, there may exist only one Apo E structural gene locus but with mutations at two sites susceptible to posttranslational modification.

Received June 4, 1979; revised October 25, 1979.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft.

[1] Institut für Humangenetik der Universität Marburg, Marburg, West Germany.

[2] Institut für Humangenetik der Universität Göttingen, West Germany.

INTRODUCTION

Apolipoprotein E is a major component of human very low density lipoproteins (VLDL) and exhibits a genetic polymorphism that is demonstrated by electrofocusing of the urea-soluble proteins from VLDL (Apo-VLDL) in the presence of 6 M urea [1–4]. Two, three, or four polymorphic forms of Apo E designated Apo E-I, Apo E-II, Apo E-III, and Apo E-IV (having isoelectric points of 5.3, 5.5, 5.6, and 5.75, respectively) may be differentiated in individual preparations [1, 2].

Population and family data show that Apo E-IV is present in only 28% of individuals from a German population sample and that this variant is under the control of two autosomal genes, *Apo E4*$^+$ and *Apo E4*^{0} [1]. The *Apo E4*$^+$ allele is dominant over *Apo E4*^{0}.

A further polymorphism of Apo E is defined by the ratio of Apo E-II/Apo E-III and is controlled by two autosomal codominant alleles, *Apo E*^{n} and *Apo E*^{d}, which determine the phenotypes Apo E-N, Apo E-ND, and Apo E-D [2, 4]. Phenotype Apo E-D is characterized by deficiency of Apo E-III.

The genes *Apo E*^{n} and *Apo E*^{d} have a significant influence on the concentration of plasma cholesterol and determine three overlapping cholesterol distributions in the population [5]. Moreover, all individuals of phenotype Apo E-D have primary dysbetalipoproteinemia (dyslipoproteinemia type III), and some have severe clinical hyperlipoproteinemia type III [1, 4, 6, 7]. The Apo E-N/D polymorphism, therefore, is possibly of high biological significance.

We show here that the *Apo E-N/D* and *Apo E4* loci are closely linked and that the Apo E polymorphism may be the first genetic system in man involving mutation at an operator/promoter region.

MATERIALS AND METHODS

Apo E phenotypes were determined in 480 blood donors from Marburg/Lahn, West Germany, and in 227 individuals with various forms of hyperlipoproteinemia by the heparin-Mg-precipitation/isoelectric focusing (IEF) procedure [1, 3]. The frequency of Apo E-N/D [1] and Apo E4 phenotypes [2] in the blood donors have been reported previously.

The segregation of Apo E phenotypes was studied in 72 matings with a total of 167 children. Pedigrees of 58 of these matings have been published in connection with our studies on the genetics of hyperlipoproteinemia type III [4]. Most probands in these kindreds were of phenotype Apo E-D.

VLDL were isolated from human sera by a standard ultracentrifugation procedure. Lipids were extracted from VLDL by acetone/ethanol 1:1 (v/v) at −20°C, and Apo E was isolated from Apo-VLDL by preparative dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) [8]. The polymorphic forms of Apo E were isolated by preparative flat bed IEF in granulated gels as outlined [9]. Analytical IEF of urea-soluble Apo-VLDL was performed in polyacrylamide gel rods containing ampholines, pH 3.5–9.0 (LKB, Bromma, Sweden).

SDS-PAGE according to Neville [10] was performed in gel slabs of 11% acrylamide monomer concentration using the Havanna electrophoresis equipment (Desaga, Heidelberg, West Germany).

STATISTICAL METHODS

Association of phenotypes was calculated according to Race and Sanger [11]. Gene frequencies were estimated according to Bernstein [12], and haplotype frequencies were calculated by the method of Mattiuz [13].

RESULTS

Frequencies of Apo E-Phenotypes

IEF of Apo-VLDL yields five different Apo E phenotypes [1, 2]. The frequency distribution of the Apo E phenotypes in blood donors and individuals with various forms of hyperlipoproteinemia is shown in table 1. It is obvious from the data that the Apo E4+ variant is less frequent in phenotype Apo E-ND (16.4%, Apo E4+) than in phenotype Apo E-N (30.1%, Apo E4+). No Apo E-D/E4(+) phenotypes were observed. Our previously published material includes 38 individuals of phenotype Apo E-D; of these, 23 were nonrelated cases and none had the E4+ variant pattern ($P < .001$). Additionally, 33 more unrelated individuals of phenotype Apo E-D have been detected in our laboratory during routine diagnosis of hyperlipoproteinemia type III. All were Apo E4 negative.

From the frequency distribution of Apo E phenotypes in blood donors, a highly significant association of the *Apo E-N/D* and *Apo E4* loci was calculated ($P \sim .01$, table 2), which may be explained by assuming very close linkage of the two *Apo E* loci and a linkage disequilibrium of $\Delta = .0147$. According to this model, Apo E phenotypes are under the control of three haplotypes: *Apo* $E^n/E4^+$, *Apo* $E^n/E4^0$, and *Apo* $E^d/E4^0$, with frequencies of .1452, .7660, and .0888, respectively (table 3). In this model, haplotype *Apo* $E^d/E4^+$ does not exist. The observed frequencies of the five phenotypes are in good agreement with those expected from Hardy-Weinberg's law. The linkage disequilibrium is in the same order as that for HLA-A and HLA-B [14].

TABLE 1

FREQUENCY OF APO E PHENOTYPES IN BLOOD DONORS AND HYPERLIPIDEMICS

	No.	Apo E-IV +	Apo E-IV −
Blood Donors:			
Apo E-N	408 (83.3%)	119 (29.2%)	289 (70.8%)
Apo E-ND	77 (15.7%)	13 (16.9%)	64 (83.1%)
Apo E-D	5 (1.0%)	0 (0%)	5 (100%)
Total	490	132 (26.9%)	358 (73.1%)
Hyperlipidemics:			
Apo E-N	166 (73.1%)	54 (32.5%)	112 (67.5%)
Apo E-ND	57 (25.1%)	9 (15.8%)	48 (84.2%)
Apo E-D	4 (1.8%)	0 (0%)	4 (100%)
Total	227	63 (27.8%)	164 (72.2%)
All:			
Apo E-N	574 (80.1%)	173 (30.1%)	401 (69.9%)
Apo E-ND	134 (18.7%)	22 (16.4%)	112 (83.6%)
Apo E-D	9 (1.2%)	0 (0%)	9 (100%)
Total	717	195 (27.2%)	522 (72.8%)

TABLE 2

ASSOCIATION OF APO E-ND AND APO E4 IN BLOOD DONORS

	Apo E-N	Apo E-ND + D	No.
Apo E4+	119	13	132
Apo E4−	289	69	358
Total	408	82	490
		$\chi^2_{df1} = 6.15$; $P \sim .01$	

Four-Haplotype Model

The frequency distribution data of Apo E phenotypes may alternatively be explained by a second model where all four haplotypes, *Apo* $E^n/E4^+$, *Apo* $E^n/E4^0$, *Apo* $E^d/E4^+$, and *Apo* $E^d/E4^0$, exist, but where the gene *Apo* $E4^+$ is not expressed when in coupling with *Apo* E^d. In the four-haplotype model, *Apo* $E^d/E4^0$ and *Apo* E^dE4^+, together, have the same frequency as *Apo* $E^d/E4^0$ in the three-haplotype model. The four-haplotype model would also explain the absence of phenotype Apo E-D/E4+ and the observation that Apo E4+ is only about half as frequent in phenotype Apo E-ND compared with phenotype Apo E-N. However, the assumption of an extreme linkage disequilibrium between both loci is not necessary in this model. The introduction of a fourth haplotype which is recessive regarding expression of E-IV yields 10 hypothetical genotypes. Since only five phenotypes can be distinguished, the same χ^2 (although with only 1 df) as in the three-haplotype model is obtained when testing Hardy-Weinberg's law (see table 3).

Segregation of Apo E-Phenotypes

The segregation of Apo E phenotypes in 72 matings with a total of 167 children is shown in table 4. As Apo E^n and Apo $E4^+$ cosegregate in all informative matings, these data also support the assumption of a close linkage between the *Apo E4* and *Apo E-N/D* loci, which are compatible with both formal genetic models, since both, the three- and four-haplotype models, would yield identical qualitative and quantitative segregation data in all mating types. Independent segregation of the genes at both the *E4* and *E-N/D* loci, however, should yield significantly different segregation data from those observed. Under the assumption of independent segregation and excluding D/4+ phenotypes which do not occur, 8.1 individuals were expected in classes which are forbidden under the haplotype hypothesis ($P = .5\%$).

Biochemistry of the Apo E Polymorphism

Crossed immunoelectrofocusing of Apo-VLDL against a monospecific anti-Apo E serum shows that the polymorphic forms E-I, E-II, and E-III of Apo E are immunochemically identical [9]. Also, the minor band occurring in the Apo E-III position of some individuals of phenotype Apo E-D reacts with the anti-Apo E serum. The Apo E-IV form of the protein also was shown to be immunochemically identical to the other Apo E proteins by double diffusion against anti-Apo E [2]; therefore, the

TABLE 3

FREQUENCY OF APO E PHENOTYPES ACCORDING TO THE THREE-HAPLOTYPE MODEL

Haplotype frequency	$E^n/E4^0 = .7660$		$E^n/E4^+ = .1452$		$E^d/E4^0 = .0888$	
Genotype	$\frac{E^n/E4^+}{E^n/E4^+}$	$\frac{E^n/E4^+}{E^n/E4^0}$	$\frac{E^n/E4^0}{E^n/E4^0}$	$\frac{E^n/E4^+}{E^d/E4^0}$	$\frac{E^n/E4^0}{E^d/E4^0}$	$\frac{E^d/E4^0}{E^d/E4^0}$
Genotype frequency	.0211	.2224	.5868	.0258	.1360	.0079
Phenotype	Apo E-N/4+		Apo E-N/4−	Apo E-ND/4+	Apo E-ND/4−	Apo E-D/4−
Phenotype frequency	.2435		.5868	.0258	.1360	.0079
Phenotype (no. expected)	119.3		287.5	12.6	66.6	3.9
Phenotype (no. observed)	119		289	13	64	5
χ^2	.0008		.0078	.0127	.1015	.3103
	$\chi^2_{df2} = .43$; $P \sim .8$					

NOTE. —Expected phenotype frequencies remain unchanged if a fourth haplotype ($E^d/E4^+$) is introduced (see text).

TABLE 4

SEGREGATION ANALYSIS OF APO E PHENOTYPES

MATING	No.	PHENOTYPES OF CHILDREN					
		N/4−	N/4+	ND/4−	ND/4+	D/4−	D/4+
N/4− × N/4−	4	6	−*	−	−	−	−
N/4− × N/4+	4	2	5	−	−	−	−
N/4− × ND/4−	19	16	−	26	−	−	−
N/4− × ND/4+	5	−	5	4	−	−	−
N/4− × D/4−	5	−	−	12	−	−	−
N/4− × D/4+	0†	...	...	...	...	...	...
N/4+ × N/4+	0	...	...	...	...	...	...
N/4+ × ND/4−	5	2	5	2	2	−	−
N/4+ × ND/4+	2	−	4	3	2	−	−
N/4+ × D/4−	4	−	−	5	6	−	−
N/4+ × D/4+	0	...	...	...	...	...	...
ND/4− × ND/4+	3	−	1	3	0	2	−
ND/4− × ND/4−	12	8	−	14	−	9	−
ND/4− × D/4−	4	−	−	3	−	8	−
ND/4− × D/4+	0	...	...	...	...	...	...
ND/4+ × ND/4+	1	−	0	−	2	0	−
ND/4+ × D/4−	3	−	−	−	5	2	−
ND/4+ × D/4+	0	...	...	...	...	...	...
D/4− × D/4−	1	−	−	−	−	3	−
D/4− × D/4+	0	...	...	...	...	...	...
D/4+ × D/4+	0	...	...	...	...	...	...
Total	72	34	20	72	17	24	0
				167			

* − = forbidden phenotypes under both haplotype models.
† 0 = not observed.

bands separated by IEF presumably represent polymorphic forms of one protein rather than nonidentical polypeptide chains. Apo E from phenotypes Apo E-N and Apo E-D do, however, differ in their apparent M_r in SDS-PAGE (fig. 1), reflecting a difference in size of Apo E-II and Apo E-III, the major polymorphic forms of the two phenotypes [9]. The apparent M_r determined for Apo E-II is ~ 34,500, and that for Apo E-III, ~ 33,000. The heterozygous phenotype Apo E-ND exhibits both Apo E species (fig. 1). The isolated Apo E-IV comigrates with Apo E-III in SDS-PAGE (not shown), and, thus, both proteins are of similar or identical molecular weight.

DISCUSSION

The four-haplotype (model A) and the three-haplotype (model B) hypotheses that are represented schematically in figure 2 cannot be distinguished by formal genetic criteria. Both the frequency of Apo E phenotypes and the segregation in families are compatible with either model. Also in both models, one additional assumption has to be made; namely, there is extreme linkage disequilibrium in the three-haplotype model and a control function of the *Apo E-N/D* locus over the *E4* locus in the four-haplotype model. However, the latter model seems much more attractive in view of the biochemical observations and may be used as a working hypothesis for further

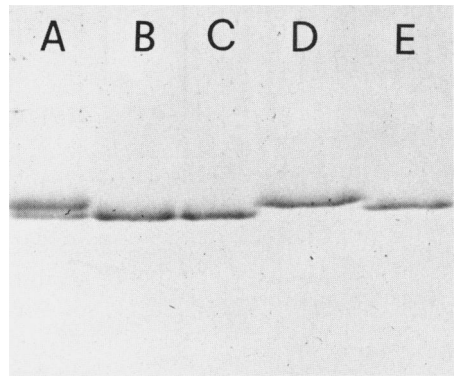

FIG. 1. —PAGE in presence of SDS of Apo E from phenotypes Apo E-N (*B*, *C*, *E*), Apo E-ND (*A*), and Apo E-D (*D*). Anode at bottom. Staining: Coomassie brilliant blue.

experiments. In contrast, the three-haplotype model does not explain any of the observed phenomena. An epistatic model that also was considered does not lead to association of the phenotypes and, hence, had to be rejected.

The quantitative character of the Apo E-N/D polymorphism in combination with the higher M_r of Apo E-II compared with Apo E-III suggest that Apo E-II is a precursor of Apo E-III and that *Apo E-N/D* is not a structural gene locus for Apo E, but, rather, is involved in the posttranslational modification of Apo E by proteolytic cleavage or deglycosilation. Hence, this locus could represent either a structural gene coding for a degrading enzyme or a gene controlling the activity of that enzyme. If one assumes that the *Apo* E^n gene is an operator/promoter region controlling a set of genes where one gene is responsible for the posttranslational cleavage of Apo E-II and a second for the absence or presence of Apo E-IV, then a mutation of that gene (*Apo* $E^n \rightarrow$ *Apo* E^d) could affect both structural loci on the same chromosome. This model, therefore, would explain both the genetic and the biochemical data.

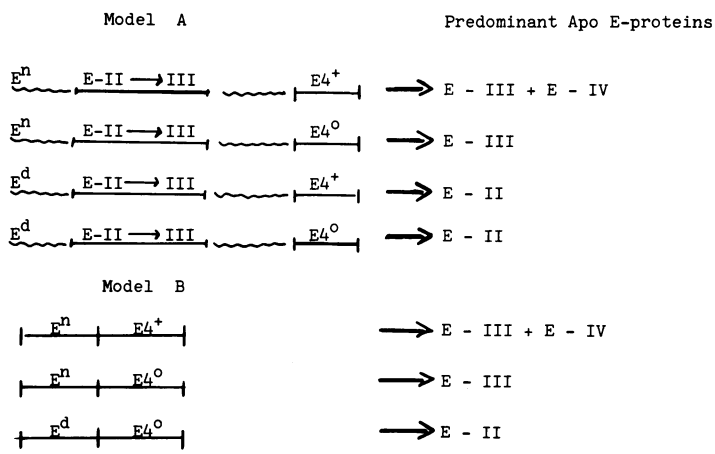

FIG. 2. —Schematic representation of the four- (*A*) and three-haplotype (*B*) models (for explanation, see text).

A third alternative interpretation might also be considered: namely, that the Apo E polymorphism is under the control of only one locus, the structural gene locus for Apo E. The primary gene product (Apo E) then may contain two sites that make it susceptible to posttranslational modification: one site recognized by an enzyme converting E-II into the E-III form (site 1), and a second, recognized by another enzyme converting E-III into the E-IV form (site 2). Mutations at these sites then would cause the protein to accumulate in the E-II (site 1) or E-III form (site 2), respectively. An additional assumption in this model, however, is that posttranslational modification at site 1 (E-II → E-III) is a prerequisite for modification at site 2 (E-III → E-IV). This one-locus model also implies a precursor-product relationship between Apo E-II and Apo E-III and explains the absence of phenotype Apo E-D/E4(+). Two mutation sites within a single *Apo E* gene would create apparent haplotypes and, hence, would explain the type of segregation data obtained. One would, however, expect that the homozygous genotype *Apo* $E^n/4^+$/*Apo* $E^n/E4^+$ would only, or at least preferentially, produce Apo E-IV and little, if any, Apo E-III. This phenotype, however, has not been observed in our material. This same problem is also unresolved in the four-haplotype model (model A).

The genetic situation in the Apo E system is complex and far from being solved. Biochemical analyses of the Apo E polymorphic forms and of the synthesis of these proteins in cell culture that are now in progress in our laboratory will help us, hopefully, to better understand this interesting system.

ACKNOWLEDGMENT

We thank Dr. Michael S. Brown for valuable discussion and his suggestion of the one-locus model.

REFERENCES

1. UTERMANN G, HEES M, STEINMETZ A: Polymorphism of apolipoprotein E and occurrence of dysbetalipoproteinemia in man. *Nature* 269:604–607, 1977
2. UTERMANN G, BEISIEGEL U, HEES M, MÜHLFELLNER G, PRUIN N, STEINMETZ A: A further genetic polymorphism of apolipoprotein E from human very low density lipoproteins, in *Protides of the Biological Fluids*, edited by PEETERS H, New York, Pergamon Press, 1977, pp 285–288
3. UTERMANN G, ALBRECHT G, STEINMETZ A: Polymorphism of apolipoprotein E. I. Methodological aspects and diagnosis of hyperlipoproteinemia type III without ultracentrifugation. *Clin Genet* 14:351–358, 1978
4. UTERMANN G, VOGELBERG KH, STEINMETZ A, ET AL.: Polymorphism of apolipoprotein E. II. Genetics of hyperlipoproteinemia type III. *Clin Genet* 15:37–62, 1979
5. UTERMANN G, PRUIN N, STEINMETZ A: Polymorphism of apolipoprotein E. III. Effect of a single polymorphic gene locus on plasma lipid levels in man. *Clin Genet* 15:63–72, 1979
6. UTERMANN G, JAESCHKE M, MENZEL J: Familial hyperlipoproteinemia type III. Deficiency of a specific apolipoprotein (Apo E-III) in the very low density lipoproteins. *FEBS Lett* 56:352–355, 1975
7. UTERMANN G, CANZLER H, HEES M, ET AL.: Studies on the metabolic defect in Broad-β disease (hyperlipoproteinemia type III). *Clin Genet* 12:139–154, 1977
8. UTERMANN G: Isolation and partial characterisation of an arginine-rich apoprotein from human plasma very low density lipoproteins: apolipoprotein E. *Hoppe Seyler's Z Physiol Chem* 356:1113–1121, 1975
9. UTERMANN G, WEBER W, BEISIEGEL U: Different mobility in SDS-polyacrylamide gel

electrophoresis of apolipoprotein E from phenotypes Apo E-N and Apo E-D. *FEBS Lett* 101:21–26, 1979
10. NEVILLE DM: Molecular weight determination of protein dodecylsulfate complexes by gel electrophoresis in a discontinuous buffer system. *J Biol Chem* 246:6328–6334, 1971
11. RACE RR, SANGER R: *Blood Groups in Man*, 5th ed. Oxford and Edinburgh, Blackwell, 1968
12. BERNSTEIN F: Fortgesetzte Untersuchungen aus der Theorie der Blutgruppen. *Z indukt Abstamm u Vererb lehre* 56:233–237, 1930
13. MATTIUZ PL, IHDE D, PIAZZA A, CEPPELLINI R, BODMER WF: New approaches to the population genetic and segregation analysis of the HLA system, in *Histocompatibility Testing*, edited by TERASAKI PI, Copenhagen, Munksgaard, 1970, pp 193–205
14. MCMICHAEL A, MCDEVITT H: The association between the HLA system and disease, in *Progress in Medical Genetics*, (N. S.), vol 2, edited by STEINBERG AG, BEARN AG, MOTULSKY AG, CHILDS B, Philadelphia-London-Toronto, WB Saunders, 1977, pp 39–100

Many laboratories now use high resolution banding techniques, so an internationally agreed system of nomenclature is now necessary for the new information available. The Standing Committee on Human Cytogenetic Nomenclature decided in Edinburgh in July 1979: (1) that a NOMENCLATURE FOR HIGH RESOLUTION BANDING OF HUMAN CHROMOSOMES is necessary, and (2) that it is essential that this nomenclature be an extension of An International System for Human Cytogenetic Nomenclature (1978)—ISCN (1978). The standing committee and its consultants are now working on this revision and hope to publish it late in 1980. We wish to draw this to the attention of human cytogeneticists who may be concerned about the lack of an agreed nomenclature to deal with the large number of bands that can now be visualized.