Am. J. Hum. Genet. 52:808–826, 1993

Polymorphic Haplotypes and Recombination Rates at the LDL Receptor Gene Locus in Subjects With and Without Familial Hypercholesterolemia Who Are from Different Populations

André R. Miserez,* Herbert Schuster,† Nicole Chiodetti,* and Ulrich Keller*

*Departments of Research and Internal Medicine, University Hospital, Basel; and †Medizinische Poliklinik der Universität, Munich

Summary

RFLPs at the low-density lipoprotein (LDL) receptor locus for *Taq*I, *Stu*I, *Hinc*II, *Ava*II, *Apa*LI (5′ and 3′), *Pvu*II, and *Nco*I were studied in Swiss and German families with familial hypercholesterolemia (FH). A total of 1,104 LDL receptor alleles were analyzed using Southern blotting and new PCR-based techniques for detection of the *Taq*I, *Stu*I, *Hinc*II, *Ava*II, *Nco*I RFLPs. Two hundred fifty-six independent haplotypes from 368 individuals of 61 unrelated Swiss families, as well as 114 independent haplotypes from 184 subjects of 25 unrelated German families, were constructed. In 76 families, clinical diagnosis of FH was confirmed by cosegregation analysis. Of the 43 unique haplotypes consisting of seven RFLPs detected in the Swiss and Germans, only 9 were common in both population samples. Analysis of linkage disequilibrium revealed nonrandom associations between several of the investigated RFLPs. *Apa*LI (5′), *Nco*I, *Pvu*II, *Taq*I, and *Ava*II or *Hinc*II were particularly informative (cumulative informativeness .85). Relative frequencies, heterozygosity indexes, and PICs of the RFLPs from the Swiss and Germans were compared with values calculated from reported haplotype data for Italians, Icelanders, North American Caucasians, South African Caucasians, and Japanese. Pairwise comparisons of population samples by common RFLPs demonstrated unexpected differences even between geographically adjacent populations (e.g., the Swiss and Germans). Furthermore, genetic distances from the Germans to the other Caucasians were larger than to the Japanese. An unexpected lack of correlation between linkage disequilibria and physical distances was detected for the German and Japanese data, possibly because of nonuniform recombination with excessively high rates between exon 13 and intron 15. Hence, the present study revealed a striking variety of polymorphic haplotypes and heterogeneity of RFLP frequencies and recombination rates among the seven population samples.

Introduction

Familial hypercholesterolemia (FH) is caused by defects of synthesis of low-density lipoprotein (LDL) receptors in the endoplasmic reticulum, by defects in their transport to the Golgi apparatus, by alterations of binding to the LDL particles or of clustering and internalization of the receptor-ligand complexes. The disease has been associated with more than 100 mutant alleles at the LDL receptor locus. FH is inherited as an autosomal dominant trait and is one of the most common inborn errors of metabolism in human beings. The population frequency of heterozygotes is approximately 1/500, and that of homozygotes is approximately 1/1 million (Brown and Goldstein 1991). However, clinical signs of FH, such as markedly elevated LDL cholesterol levels, tendon xanthomas, and atheromas, are not always present. Furthermore, a similar clinical syndrome is caused by a recently described disorder, familial defective apolipoprotein B-100 (FDB; Soria et al. 1989). Thus, either biochemical testing of the LDL receptor function or molecular genetic characterization of the LDL receptor gene is necessary for confirming FH.

Since the molecular defects of FH in central Europe are not known, direct DNA diagnosis as described in other populations—such as South African Afrikaners (Kotze et al. 1986, 1989*a*, 1989*c*; Brink et al. 1987; Leitersdorf et al. 1989*b*), French-Canadians (Hobbs et al. 1987*a*; Ma et al. 1989), Lebanese (Lehrman et al. 1987*b*), and Finns (Aalto-Setälä et al. 1989)—is not fea-

Received April 28, 1992; final revision received December 1, 1992.

Address for correspondence and reprints: Dr. A. R. Miserez, Departments of Research and Internal Medicine, University Hospital, Petersgraben 4, CH-4031 Basel, Switzerland.

0002-9297/93/5204-0022$02.00

sible. However, if it is assumed that there is no recombination between the generations, the RFLP method (Botstein et al. 1980) allows unequivocal confirmation of the disease, provided that sufficient family members and RFLPs are analyzed (Humphries et al. 1985, 1989; Kotze et al. 1986, 1989*a*, 1989*b*, 1989*c*; Leitersdorf et al. 1989*b*; Schuster et. al. 1989; Taylor et al. 1989; Daga et al. 1990; Yamakawa et al. 1991). FH can thereby be diagnosed even in cases in which clinical signs are not present (e.g., in children). Furthermore, it is possible to investigate the heterogeneity of haplotypes corresponding to underlying genetic defects in a given population or to detect characteristic haplotypes and possible founder gene effects (Brink et al. 1987).

Objectives of the present study were to evaluate the use of eight RFLPs for confirmation of FH in Swiss and Germans and to study the heterogeneity and frequency distribution of the different haplotypes associated with defective and normal (i.e., not defective) LDL receptor alleles in these two populations. Since the LDL receptor locus has now been characterized in several populations by numerous different RFLPs, our data were compared with those from other population samples to assess variability of haplotypes and mutations leading to FH and to obtain new insights into the structure of the gene.

Subjects and Methods

Subjects

Five hundred fifty-two individuals, 368 from 61 unrelated Swiss and 184 from 25 unrelated German families, were studied. Individuals suggested not to be autochthonous, because of possible immigration of ancestors or from analysis of surnames and maiden names of grandparents, were excluded. Among the 61 Swiss families, 35 were from northern, 5 from southwestern, 5 from western, and 16 from eastern parts of Switzerland. In these 61 Swiss families, there were 186 subjects (61 unrelated index patients and 125 relatives) with clinically suggested FH and 182 individuals (including 77 unrelated spouses) without hypercholesterolemia. The 25 German families from the Munich area consisted of 109 individuals (25 unrelated index patients and 84 relatives) with suggested FH and 75 normocholesterolemic individuals (including 32 unrelated spouses). FH was clinically diagnosed in the presence of LDL cholesterol levels exceeding the 95th percentile corrected for age and sex, in at least three family members of two generations. Major rearrangements at the LDL receptor locus in the 61 Swiss and 25 German families were excluded by restriction analysis according to the methods of Hobbs et al. (1988) and Humphries et al. (1989). In addition, familial defective apolipoprotein B-100 (FDB) due to the 3500 G-to-A mutation was excluded in all families by a one-step PCR (Schuster et al. 1992). All subjects gave their informed consent to participate in this study.

Methods

Lipoprotein analysis.—Blood samples were drawn 12 h postprandially. LDL cholesterol was determined after precipitation with heparin (Merck, Darmstadt), and HDL cholesterol was determined after precipitation of the apolipoprotein B–containing particles (Boehringer Mannheim, Germany) on a Hitachi analyzer model 737.

RFLP analysis.—The *Apa*LI (5′), *Pvu*II, and *Apa*LI (3′) RFLPs were detected using Southern blotting according to the methods of Humphries et al. (1985) and Leitersdorf and Hobbs (1987), with modifications: after extraction from white blood cells by the salting-out method (Miller et al. 1988), total genomic DNA was digested with restriction enzymes, electrophoresed, and transferred to nylon membranes by capillary blotting. Membranes were hybridized with a [α-^{32}P]dCTP-labeled 3′ cDNA probe (pLDLR-2HH1) containing a 1.9-kb fragment of the full-length cDNA which hybridizes with exons 11–18 of the LDL receptor gene (Yamamoto et al. 1984) and were exposed for 24–72 h. After PCR amplification and digestion, the *Taq*I, *Stu*I, *Hin*cII, *Ava*II, and *Nco*I RFLPs, as described by Yamakawa et al. (1987), Kotze et al. (1986), Leitersdorf and Hobbs (1988), Hobbs et al. (1987*b*), and Kotze et al. (1987), respectively, were directly detected on agarose and polyacrylamide gels by staining with ethidium bromide. Intron 4 containing the *Taq*I RFLP was amplified using the primers T1 (5′ ATCTGACGAGGAAAACTGCG 3′) and T2 (5′ CAGGGCGACAGGTGGCCACAG 3′), exon 13 containing the *Ava*II RFLP was amplified using the primers A1 (5′ GACAAAGTATTTTGGACAG 3′) and A2 (5′ CTCTTGGCTGGGTGAGGTTG 3′), and the *Nco*I RFLP–containing part of exon 18 was amplified using primers N1 (5′ CAATCTTGTCGTTGATGG 3′) and N2 (5′ CAAACGATCCAGACTGGAGG 3′). The *Stu*I and *Hin*cII RFLPs were investigated using modified methods of Leitersdorf et al. (1989*a*) and Leitersdorf and Hobbs (1988), respectively.

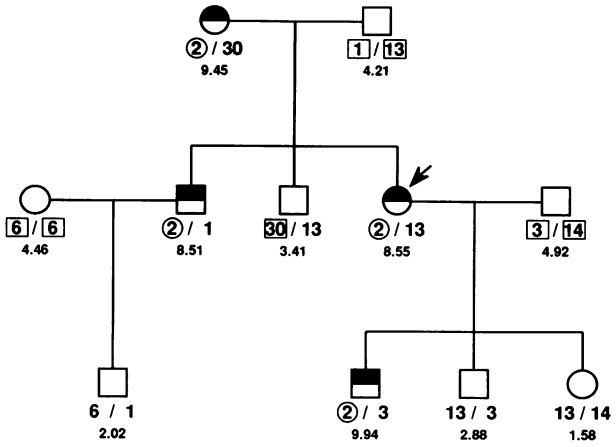


Figure 1 Swiss family informative for eight RFLPs. One of the two alleles of the index patient (marked by the arrow) (haplotype 2; see table 1) cosegregated with the phenotype of FH, characterized by elevated LDL cholesterol (given in mmol/liter, below the haplotypes), since the combination of this haplotype and the FH phenotype was also detected in three other family members (in one of the brothers of the index patient, in one of her children, and in her mother). However, haplotype 2 was not observed in the clinically unaffected family members. Three of the individuals in this pedigree were normocholesterolemic spouses and not related to FH patients of the same generation. From these three spouses and from the unaffected brother of the index patient, seven independent haplotypes were detected (marked by squares). Thus, a haplotype of a probably defective LDL receptor gene (haplotype 2) was characterized by eight RFLPs, and the clinical diagnosis of FH was confirmed by cosegregation analysis. Circles denote defective alleles; and squares denote independent normal alleles.

Haplotype analysis.—Haplotypes for each subject were constructed by segregation analysis, on the basis of the assumption that there had been no recombinational events at the LDL receptor locus within a given family. From each pedigree only independent haplotypes (i.e., the FH haplotype and the haplotypes of the normocholesterolemic spouses and normocholesterolemic relatives of the FH patients; fig. 1) were included in further analyses. Haplotypes for defective and normal LDL receptor alleles were determined on the bases of cosegregation between particular haplotypes and normo- or hypercholesterolemia in sibs or offsprings.

Statistical methods.—Physical distances were calculated on the bases of the human LDL receptor genome map (Yamamoto et al. 1984) and intron lengths given by Südhof et al. (1985). The allele frequency of each individual RFLP was estimated by gene counting of the independent chromosomes. Tests for goodness of fit to Hardy-Weinberg equilibrium, frequencies (including standard errors [SEs]) of presence of the restriction sites, and heterozygosity indexes were calculated using standard methods (e.g., see Chakravarti et al. 1984). PIC values at a particular RFLP were computed according to the algorithm of Botstein et al. (1980). Published haplotypes from Italians (Daga et al. 1990), Icelanders (Taylor et al. 1989), North American Caucasians (Leitersdorf et al. 1989*a*), South African Caucasians (Kotze et al. 1989*a*), and Japanese (Yamakawa et al. 1991), as well as the haplotypes from the Swiss and Germans, were used to estimate frequencies of the restriction sites, heterozygosities, and PIC values in the respective population samples. Heterogeneity of frequencies of cleavage at a particular RFLP among the seven populations were assessed in a univariate approach using Fisher's exact test (SAS Software, Cary, NC). Haplotype data from these populations were also used for (1) direct comparison by construction of new haplotypes on the basis of the common RFLPs of two population samples in each case and (2) computation of genetic distances. Haplotype frequency differences between FH patients and controls, as well as those between different population samples, were assessed by χ^2 tests according to Cochran's criterion: rows with expected frequencies not exceeding 1 were combined, as were rows in which expected frequencies did not exceed 5, until at least 80% of the expected frequencies met the criterion (e.g., see Bland 1989). Genetic distances d_{st} (and their SEs) between different populations s and t were determined using the algorithms of Cavalli-Sforza and Bodmer (1971, pp. 704–708) and Ingram and Bloch (1984, pp. 179–184), given as

$$d_{st} = \frac{\sum_{j=1}^{M} (K_j - 1)d_{stj}}{\left(\sum_{j=1}^{M} K_j\right) - M},$$

with SE

$$SE(d_{st}) = \sqrt{\frac{\sum_{j=1}^{M} d_{stj}^2 - \left(\sum_{j=1}^{M} d_{stj}\right)^2 \Big/ M}{(M)(M-1)}},$$

where N is the number of populations, M is the number of marker loci studied, K_j is the number of alleles at marker j for $j = 1, \ldots, M$, p_{ijs} is the frequency of allele i at marker j in population s ($s = 1, \ldots, N$; $j = 1, \ldots, M$; and $i = 1, \ldots, K_j$),

$$d_{stj} = \frac{4(1 - \cos\theta_{stj})}{K_j - 1}, \quad \text{and} \quad \cos\theta_{stj} = \sum_{i=1}^{K_j} \sqrt{p_{ijs} p_{ijt}}.$$

Linkage disequilibrium statistics were studied by calculating the correlation coefficient Δ (Hill and Robertson 1968) and the relative linkage disequilibrium D' (Lewontin 1964) between the polymorphic sites at two loci. The null hypothesis of no association between RFLPs was tested using Fisher's exact test with fixed marginal totals (Chakravarti et al. 1984):

$$P = \sum \frac{(n_1 + n_2)!(n_3 + n_4)!(n_1 + n_3)!(n_4 + n_2)!}{n! n_1! n_2! n_3! n_4!}.$$

Probability was also estimated using the Bonferroni correction as described, by Leitersdorf et al. (1989*a*), for Δ values. The number of significant Δ values expected by chance alone was computed as $k_{exp} = m\alpha$ (significance level $\alpha = .05$) $\pm\sqrt{m\alpha(1 - \alpha)}$. The probability of obtaining k or more significant Δ values among m tests under the null hypothesis of no association was (Chakravarti et al. 1984)

$$P = \sum_{r=k}^{m} \binom{m}{r} \alpha^r (1 - \alpha)^{m-r}.$$

Cumulative informativeness of RFLPs in combination was calculated according to the method of Chakravarti and Buetow (1985) for autosomal dominant diseases:

$$I_{A\,cum} = \left[1 - \left(\sum_{i=1}^{k} a_i b_i\right)\right] - \left(\sum_{i=1}^{k} a_i b_i \sum_{i} a_i^2\right) + \sum_{i} a_i^3 b_i.$$

An approximation of the relationship between linkage disequilibria $|\Delta|$ and physical distances d (in kb) for small distances and estimations of $4N_e k$ values were assessed using the formula of Hill and Robertson (1968), Ohta and Kimura (1969), and Sved (1971): $\Delta^2 \approx 1/(1 + 4N_e kd)$, where N_e is the effective population size and k is the recombination frequency per unit length (in kb). To bypass the problem of a lack of independence of the $|\Delta|$ values, Mantel's matrix correspondence method was applied (Mantel 1967; Chakraborty et al. 1987). Significance levels of the correlations between $|\Delta|$ and d were evaluated by calculation of large numbers of random permutations after development of a program using Fortran subroutines from IMSL (Houston). These methods were also used to assess relationships between $|\Delta|$ and d, as well as $4N_e k$ values from published Δ values for Italians (Daga et al. 1990), North Americans (Leitersdorf et al. 1989*a*), and Japanese (Yamakawa et al. 1991). All calculations were performed on Macintosh and DEC computers.

Results

Cosegregation Analysis in the Swiss and German Families

A total of 1,104 LDL receptor alleles (736 from 368 members of 61 unrelated Swiss families and 368 from 184 individuals of 25 unrelated German families) were analyzed for eight RFLPs, with the restriction enzymes *Taq*I, *Stu*I, *Hin*cII, *Ava*II, *Apa*LI (5′), *Pvu*II, *Nco*I, and *Apa*LI (3′). The RFLPs extended from intron 4 to the noncoding region, 14.6 kb downstream to the 3′ end of exon 18 of the LDL receptor gene, and span a total distance of approximately 43.3 kb. In the Swiss sample, independent haplotypes consisting of eight RFLPs were constructed for 256 chromosomes: 52 FH haplotypes (one patient was a compound heterozygote) and 204 control haplotypes. In the Germans, seven RFLPs were used to construct independent haplotypes for 114 chromosomes: 25 FH haplotypes and 89 control haplotypes. χ^2 tests revealed neither in the Swiss controls nor in the German controls statistically significant departures from Hardy-Weinberg equilibrium. Of $2^8 = 256$ possible combinations for eight RFLPs, a total of 37 unique haplotypes from 178 Swiss individuals were detected (table 1). The frequency of the individual haplotypes ranged from 1/204 to 50/204 (24.5%) in controls and from 1/52 to 12/52 (23.1%) in FH patients. Haplotype 4 in the Swiss was more than twice as frequent in FH patients than in controls. However, no statistically significant higher relative frequency of a particular FH haplotype compared with the corresponding control haplotype was observed. In the German population sample, of $2^7 = 128$ theoretically possible haplotypes, a total of 26 different combinations from 82 subjects were detected (table 2). The frequency of the individual haplotypes ranged from 1/89 to 22/89 (24.7%) in controls and from 1/25 to 8/25 (32.0%) in FH patients. Haplotype frequencies of German FH patients and those of controls were not significantly different. The most common haplotype in the Swiss accounted for approximately 25% of the normal chromosomes, as did that in the Germans. However, these two haplotypes differed from one another in three of the seven common RFLPs. In 51 of the 61

Table 1

Swiss Haplotypes: 256 Independent Haplotypes (51 Unrelated FH Patients and 127 Unrelated Control Subjects) Identified from Analysis of 61 Swiss Pedigrees

Haplotype	Status[a]								No. (%)	
	*Taq*I	*Stu*I	*Hinc*II	*Ava*II	*Apa*LI (5′)	*Pvu*II	*Nco*I	*Apa*LI (3′)	Control Haplotypes	FH Haplotypes
1	−	+	−	+	+	−	+	+	50 (24.51)	12 (23.08)
2	−	+	+	−	−	+	+	+	28 (13.73)	9 (17.31)
3	−	+	+	−	−	−	−	−	25 (12.25)	7 (13.46)
4	+	+	+	−	+	−	+	+	12 (5.88)	7 (13.46)
5	−	+	−	+	+	+	+	+	9 (4.41)	0
6	−	+	+	−	+	−	+	+	8 (3.92)	3 (5.77)
7	−	+	−	+	+	−	−	+	8 (3.92)	0
8	−	+	+	−	−	−	+	+	7 (3.43)	2 (3.85)
9	−	+	+	−	+	+	+	+	7 (3.43)	1 (1.92)
10	−	+	+	−	−	−	−	+	5 (2.45)	1 (1.92)
11	−	−	+	−	−	+	+	+	4 (1.96)	1 (1.92)
12	−	+	−	+	−	−	+	+	4 (1.96)	0
13	−	+	+	−	+	−	−	+	4 (1.96)	0
14	+	+	+	−	−	−	−	−	3 (1.47)	1 (1.92)
15	−	+	−	+	−	−	−	−	3 (1.47)	1 (1.92)
16	+	+	+	−	+	+	+	+	3 (1.47)	0
17	−	+	−	−	+	−	+	+	2 (.98)	1 (1.92)
18	+	+	−	+	+	−	+	+	2 (.98)	1 (1.92)
19	−	+	−	+	−	+	+	+	2 (.98)	0
20	−	+	−	+	−	+	−	−	2 (.98)	0
21	−	−	+	−	+	+	+	+	2 (.98)	0
22	−	+	+	−	−	+	+	−	2 (.98)	0
23	−	−	+	−	−	−	+	+	1 (.49)	2 (3.85)
24	−	+	+	+	+	−	+	+	1 (.49)	0
25	−	+	−	+	+	−	+	−	1 (.49)	0
26	+	+	−	−	+	−	+	+	1 (.49)	0
27	+	+	−	+	+	−	−	+	1 (.49)	0
28	−	+	+	−	−	−	+	−	1 (.49)	0
29	−	+	+	−	−	+	−	+	1 (.49)	0
30	−	+	−	+	−	−	−	+	1 (.49)	0
31	−	+	+	−	+	−	−	−	1 (.49)	0
32	−	−	−	+	+	−	+	+	1 (.49)	0
33	−	+	−	+	+	+	−	+	1 (.49)	0
34	−	+	+	+	−	−	−	−	1 (.49)	0
35	−	+	+	−	−	+	−	−	0 (0)	1 (1.92)
36	−	+	−	+	+	−	−	−	0 (0)	1 (1.92)
37	−	+	+	+	−	+	−	+	0 (0)	1 (1.92)
Total ...									204 (100[b])	52 (100[b])

[a] A plus sign denotes presence, and a minus sign denotes absence.
[b] Because of rounding error, entries do not sum to exactly 100%.

Swiss families with clinical diagnosis of FH, cosegregation between the FH phenotypes and a particular haplotype in a given family could be demonstrated (fig 1). Thus, in these 51 families, FH was molecular genetically confirmed using linked markers; in 5 of the 51 Swiss families it also was confirmed by LDL binding assays. In the remaining 10 families, FH was not diagnosed, either because noncosegregation between phenotypes and haplotypes has been observed (three families) or because parents and offspring were all

Table 2

German Haplotypes: 114 Independent Haplotypes (25 Unrelated FH Patients and 57 Unrelated Control Subjects) Identified from Analysis of 25 German Pedigrees

Haplotype	Status							No. (%)	
	*Taq*I	*Stu*I	*Hin*cII	*Ava*II	*Apa*LI (5′)	*Pvu*II	*Nco*I	Control Haplotypes	FH Haplotypes
1	−	+	+	−	−	−	−	22 (24.72)	4 (16.00)
2	+	+	−	−	+	−	+	18 (20.22)	8 (32.00)
3	−	+	+	−	−	+	+	15 (16.85)	3 (12.00)
4	+	+	−	−	+	−	−	6 (6.74)	0
5	−	+	+	−	−	−	+	4 (4.49)	0
6	−	−	+	−	−	+	+	4 (4.49)	0
7	−	+	−	−	−	−	−	2 (2.25)	2 (8.00)
8	−	+	+	+	−	−	−	2 (2.25)	0
9	+	+	+	−	−	+	+	1 (1.12)	2 (8.00)
10	−	+	+	+	+	−	+	1 (1.12)	1 (4.00)
11	−	+	−	−	−	+	+	1 (1.12)	1 (4.00)
12	+	+	−	−	−	−	+	1 (1.12)	1 (4.00)
13	−	+	−	+	+	−	+	1 (1.12)	0
14	−	+	−	−	+	−	+	1 (1.12)	0
15	−	+	+	−	−	+	−	1 (1.12)	0
16	−	−	+	−	+	+	−	1 (1.12)	0
17	−	−	+	−	+	−	−	1 (1.12)	0
18	+	+	−	−	−	+	+	1 (1.12)	0
19	+	−	+	−	−	+	+	1 (1.12)	0
20	−	+	+	+	−	+	+	1 (1.12)	0
21	−	+	+	+	−	−	+	1 (1.12)	0
22	−	−	+	−	−	+	−	1 (1.12)	0
23	+	−	−	−	−	−	+	1 (1.12)	0
24	+	+	−	−	+	+	+	1 (1.12)	0
25	−	+	+	−	+	−	+	0	2 (8.00)
26	+	+	+	−	+	−	+	0	1 (4.00)
Total ...								89 (100)	25 (100)

NOTE.—See footnotes to table 1.

heterozygous and other family members did not allow a resolution of the linkage phase (three families). Another reason to exclude families were G-to-A mutations at the codon for amino acid 3500 of the apolipoprotein B-100 gene (four families). In all of the 25 investigated German families, cosegregation between FH phenotypes and a particular haplotype was detected, and thus clinical diagnosis of FH was confirmed; in 17 of the 25 German families it was also confirmed by LDL binding assays (Spengel et al. 1982).

Assessment of Genetic Differences between Populations

Differences of characteristics of polymorphisms.—Frequencies of presence of restriction sites, heterozygosities, and PIC values of each individual RFLP in controls and FH patients from Switzerland and Germany were compared with values calculated from haplotype data of five other populations (table 3). A significant heterogeneity of frequencies of sites among the populations was revealed for six (i.e., *Taq*I, *Ava*II, *Apa*LI [5′], *Pvu*II, *Nco*I, and *Apa*LI [3′]) of the eight RFLPs. Comparison between controls and FH patients revealed, in the Swiss and Germans, none of the statistically significant differences that have been reported from other populations. Calculation of heterozygosity indexes showed that the *Hin*cII, *Apa*LI (5′), and *Nco*I RFLPs were fairly polymorphic in all populations investigated, ranging from .398 (for *Nco*I in Swiss controls) to .500 (for *Apa*LI [5′] in North Americans). Particularly high means of PIC values from controls were revealed for

Table 3

Comparisons of Relative Frequencies of Presence of the Restriction Sites and of Heterozygosity Indexes, and PIC Values, in Seven Populations

RFLP and Group	Relative Frequency ± SEM of Site							
	Swiss[a]	Germans[b]	Italians[c]	Icelanders[d]	North Americans[e]	South Africans[f]	Japanese[g]	P^h
*Taq*I:								
Controls	.108 ± .022	.337 ± .050	...	...	...	.036 ± .018*	.512 ± .045	.0001
FH patients	.173 ± .052	.480 ± .100	...	...	...	0*	...	.0001
*Stu*I:								
Controls	.961 ± .014	.899 ± .032	...	...	.943 ± .021	.946 ± .021	...	.234
FH patients	.942 ± .032	1	...	...	...	.899 ± .030	...	.254
*Hin*cII:								
Controls	.569 ± .035	.629 ± .051	...	...	...	...	...	.368
FH patients	.692 ± .064	.520 ± .100	...	...	...	...	...	.206
*Ava*II:								
Controls	.426 ± .035	.067 ± .027	.387 ± .030	...	.431 ± .045	.378 ± .046*	.149 ± .032	.0001
FH patients	.308 ± .064	.040 ± .039	.439 ± .066	...	...	.020 ± .014*	...	.0001
*Apa*LI (5′):								
Controls	.559 ± .035	.337 ± .050	.426 ± .031	.563 ± .072	.512 ± .045	.414 ± .047*	.430 ± .045	.004
FH patients	.500 ± .069	.480 ± .100	.474 ± .066	.769 ± .117	...	.020 ± .014*	...	.0001
*Pvu*II:								
Controls	.299 ± .032	.315 ± .049	.371 ± .030*	.229 ± .061	.244 ± .039	.297 ± .043	.074 ± .024	.0001
FH patients	.250 ± .060	.240 ± .085	.158 ± .048*	.077 ± .074	...	.121 ± .033	...	.218
*Nco*I:								
Controls	.725 ± .031	.596 ± .052	...	.688 ± .067	.724 ± .040	.667 ± .045*	.545 ± .045	.010
FH patients	.750 ± .060	.760 ± .085	...	.769 ± .117	...	.141 ± .035*	...	.0001
*Apa*LI (3′):								
Controls	.809 ± .028	...	.797 ± .025*	...	.732 ± .040	.730 ± .042*	.529 ± .045	.0001
FH patients	.788 ± .057	...	.667 ± .062*	...	...	.131 ± .034*	...	.0001

	Heterozygosity Index							
	Swiss[a]	Germans[b]	Italians[c]	Icelanders[d]	North Americans[e]	South Africans[f]	Japanese[g]	Mean ± SEM
*Taq*I:								
Controls	.192	.447	...	...	...	.069	.500	.302 ± .103
FH patients	.286	.499	...	...	...	.000	...	.262 ± .145
*Stu*I:								
Controls	.075	.182	...	...	.107	.102	...	.117 ± .023
FH patients	.109	.000	...	...	...	.182	...	.097 ± .053
*Hin*cII:								
Controls	.491	.467	...	...	...	...	...	.479 ± .012
FH patients	.426	.499	...	...	...	...	...	.463 ± .037

*Ava*II:								
Controls	.489	.125	.474	...	.490	.470	.254	.384 ± .064
FH patients	.426	.077	.493	...	...	.039	...	.259 ± .117
*Apa*LI (5′):								
Controls	.493	.447	.489	.492	.500	.485	.490	.485 ± .007
FH patients	.500	.499	.499	.355	...	.039	...	.378 ± .089
*Pvu*II:								
Controls	.419	.432	.467	.353	.369	.418	.137	.371 ± .042
FH patients	.375	.365	.266	.142	...	.213	...	.272 ± .045
*Nco*I:								
Controls	.398	.482	...	.429	.400	.444	.496	.441 ± .017
FH patients	.375	.365	...	.355	...	242	...	.334 ± .031
*Apa*L (3′):								
Controls	.309	...	.324	...	.393	.394	.498	.384 ± .033
FH patients	.334	...	.444	...	...	.228	...	.335 ± .062

	PIC							
	Swiss[a]	Germans[b]	Italians[c]	Icelanders[d]	North Americans[e]	South Africans[f]	Japanese[g]	
*Taq*I:								
Controls	.174	.347	...	...	...	.067	.375	.241 ± .073
*Stu*I:								
Controls	.073	.165	...	...	.102	.097	...	.109 ± .020
*Hin*cII:								
Controls	.370	.358	...	...	...	...	...	.364 ± .006
*Ava*II:								
Controls	.370	.117	.362	...	.370	.360	.221	.300 ± .043
*Apa*LI (5′):								
Controls	.372	.347	.380	.371	.375	.368	.370	.369 ± .004
*Pvu*II:								
Controls	.331	.338	.358	.291	.301	.330	.128	.297 ± .029
*Nco*I:								
Controls	.319	.366	...	.337	.320	.346	.373	.343 ± .009
*Apa*LI (3′):								
Controls	.262	...	.271	...	.316	.317	.374	.308 ± .020

[a] Data are from present study of 52 FH haplotypes and 204 control haplotypes.
[b] Data are from present study of 25 FH haplotypes and 89 control haplotypes.
[c] Data are from Daga et al. (1990) study of 57 FH haplotypes and 256 control haplotypes.
[d] Data are from Taylor et al. (1989) study of 13 FH haplotypes and 48 control haplotypes.
[e] Data are from Leitersdorf et al. (1989*a*) study of 123 control haplotypes.
[f] Data are from Kotze et al. (1989*a*) study of 99 FH haplotypes and 111 control haplotypes.
[g] Data are from Yamakawa et al. (1991) study of 121 control haplotypes.
[h] By Fisher's exact test.
* $P < .05$ (significant difference between FH haplotypes and control haplotypes in this population).

*Apa*LI (5′), *Nco*I, and *Hin*cII, each of which had individual PIC values higher than .3 in all of the investigated population samples.

Differences of haplotype frequencies.—Frequencies of haplotypes constructed from the common RFLPs of two population samples were compared to detect haplotypes which are characteristic of a given population sample. Tables were prepared for each of the 21 pairwise comparisons of the seven populations studied, and differences of haplotype frequencies were statistically assessed. For illustration, two of the comparisons are presented in tables 4 and 5: 19 of the 28 Swiss haplotypes (constructed using the seven RFLPs studied in Swiss as well as in Germans) were exclusive to the Swiss subjects. On the other hand, 15 of the 24 German haplotypes were detected solely in the Germans. Nine of the 43 haplotypes were observed in both samples. Between Swiss and Germans, the difference in frequencies of corresponding haplotypes was statistically significant (table 6). Ten of the 25 Swiss haplotypes (constructed using the six common RFLPs in Swiss and North Americans) were not detected in the North American sample, but, conversely, all of the 15 American haplotypes were present also in the Swiss (table 5). Frequencies of corresponding haplotypes were similar in the two samples: the cumulative frequency of the Swiss haplotypes 1, 18, and 24 was 53/204 (26.0%), and that of the corresponding American haplotypes 1, 9, 10, and 22 (numbered as in the original publication) was 41/123 (33.3%). The Swiss haplotype 2 was found in 28/204 (13.7%), and the respective American haplotypes 2, 12, 13, and 29 were found in 18/123 (14.6%). Overall, the frequency difference between the Swiss haplotypes and the corresponding American haplotypes was not significant.

Genetic distances.—Genetic distances at the LDL receptor locus, as parameters of genetic differences among the seven population samples, were relatively large ($d_{st} > .1$) between Germans and Swiss, Italians, North Americans, or South Africans but not between Germans and Japanese (table 6). Distances between Icelanders and the other Caucasian samples were relatively small ($d_{st} < .05$). Genetic distances between Swiss, Italian, North American, and South African Caucasians were also small ($d_{st} < .02$), particularly the distances between North Americans and Swiss or South Africans and between Italians and South Africans ($d_{st} < .01$).

Nonrandom Associations between a Pair of Loci

Under the null hypothesis of no association, RFLPs are expected to segregate independently. Since this may

Table 4

Comparison of Haplotypes from Swiss and German Controls

Swiss		Germans	
Haplotype(s)	No. (%)	Haplotype	No. (%)
1 and 25	51 (25.00)	13	1 (1.12)
2 and 22	30 (14.71)	3	15 (16.85)
3 and 10	30 (14.71)	1	22 (24.72)
4	12 (5.88)	26[a]	0
5	9 (4.41)		0
6	8 (3.92)	25[a]	0
7 and 36	8 (3.92)		0
8 and 28	8 (3.92)	5	4 (4.49)
9	7 (3.43)		0
11	4 (1.96)	6	4 (4.49)
12	4 (1.96)		0
13 and 31	5 (2.45)		0
14	3 (1.47)		0
15 and 30	4 (1.96)		0
16	3 (1.47)		0
17	2 (.98)	14	1 (1.12)
18	2 (.98)		0
19	2 (.98)		0
20	2 (.98)		0
21	2 (.98)		0
23	1 (.49)		0
24	1 (.49)	10	1 (1.12)
26	1 (.49)	2	18 (20.22)
27	1 (.49)		0
29 and 35[a]	1 (.49)	15	1 (1.12)
32	1 (.49)		0
33	1 (.49)		0
34	1 (.49)		0
37[a]	0		0
	0	4	6 (6.74)
	0	7	2 (2.25)
	0	8	2 (2.25)
	0	9	1 (1.12)
	0	11	1 (1.12)
	0	12	1 (1.12)
	0	16	1 (1.12)
	0	17	1 (1.12)
	0	18	1 (1.12)
	0	19	1 (1.12)
	0	20	1 (1.12)
	0	21	1 (1.12)
	0	22	1 (1.12)
	0	23	1 (1.12)
	0	24	1 (1.12)
Total	204 (100[b])		89 (100[b])

[a] Found only in FH patients and not in controls.

[b] Because of rounding error, entries do not sum to exactly 100%.

Table 5

Comparison of Haplotypes from Swiss and North American Controls

Swiss		North Americans	
Haplotype(s)	No. (%)	Haplotype(s)	No. (%)
1, 18, and 24	53 (25.98)	1, 9, 10, and 22	41 (33.33)
2	28 (13.73)	2, 12, 13, and 29	18 (14.63)
3 and 14	28 (13.73)	3, 11, and 15	26 (21.14)
4, 6, 17, and 26	23 (11.27)	8, 19, 21, 24, 25, and 31	14 (11.38)
5	9 (4.41)	18	1 (.81)
7 and 27	9 (4.41)	23 and 30	3 (2.44)
8	7 (3.43)	27	1 (.81)
9 and 16	10 (4.90)	20	1 (.81)
10	5 (2.45)	7	1 (.81)
11	4 (1.96)	4 and 16	7 (5.69)
12	4 (1.96)		0
13	4 (1.96)		0
15 and 34	4 (1.96)	5	3 (2.44)
19	2 (.98)	14 and 17	2 (1.63)
20	2 (.98)		0
21	2 (.98)		0
22	2 (.98)		0
23	1 (.49)		0
25	1 (.49)	6	3 (2.44)
28	1 (.49)	26	1 (.81)
29	1 (.49)	28	1 (.81)
30	1 (.49)		0
31	1 (.49)		0
32	1 (.49)		0
33	1 (.49)		0
35[a]	0		0
36[a]	0		0
37[a]	0		0
	204 (100%[b])		123 (100%[b])

[a] Found only in FH patients and not in controls.
[b] Because of rounding error, entries do not sum to exactly 100%.

not be the case, the lack of independence between two marker loci in each case was studied among 204 Swiss and 89 German chromosomes of the control subjects (table 7). Twenty of 28 linkage disequilibria (Δ) in the Swiss population sample, as well as 11 of 21 Δ values in the Germans, were significant when Fisher's exact probability was calculated. In a more conservative approach, using the Bonferroni correction, the lower limits to achieve significance for Δ values in the two population samples were $|\Delta|_{\text{Swiss}} = \sqrt{(9.758/204)} = .219$ and $|\Delta|_{\text{Germans}} = \sqrt{(9.234/89)} = .322$. Thus, 12 of 28 Δ values in the Swiss and 7 of 21 Δ values in the German sample were significant when this approximation was used. However, under the null hypothesis of no association between RFLPs, the number of values expected to be significant was $k_{exp} = 1.40$ (SE ± 1.15) for 28 tests in the Swiss and was $k_{exp} = 1.05$ (SE ± 1.00) for 21 tests in the German sample. On the other hand, the probabilities of finding, by chance alone, 12 or more significant Δ values among 28 tests and 7 or more significant Δ values among 21 tests were very low ($P = 3.493 \times 10^{-9}$ and $P = 4.871 \times 10^{-5}$, respectively). Hence, considerable linkage disequilibria existed in both population samples.

Cumulative Maximum of Informativeness

To determine the rank order in which the RFLPs should be chosen, algorithms were computed to maximize informativeness when multiple linked markers are used (table 8). The contribution to cumulative informa-

Table 6

Direct Comparison of Haplotypes from Population Samples, and Genetic Distances

	Swiss	Germans	Italians	Icelanders	North Americans	South Africans	Japanese
Swiss		70.075 7 .0001[a]	63.140 9 .0001[a]	6.806 4 .1465	13.412 7 .0627	38.216 7 .0001[a]	81.714 10 .0001[a]
Germans	1050 .0532		61.427 5 .0001[a]	12.111 4 .0165[a]	37.334 6 .0001[a]	49.003 5 .0001[a]	33.815 6 .0001[a]
Italians	.0127 .0080	.1176 .1058		4.031 2 .1333	10.648 5 .0588	13.141 6 .0409[a]	119.257 6 .0001[a]
Icelanders	.0053 .0038	.0473 .0287	.0431 .0054		.909 3 .8232	2.290 3 .5145	18.432 4 .001[a]
North Americans	.0054 .0026	.1064 .0754	.0172 .0074	.0029 .0013		11.505 6 .0740	55.743 5 .0001[a]
South Africans	.0167 .0068	.1197 .0696	.0063 .0035	.0190 .0129	.0067 .0029		84.641 6 .0001[a]
Japanese	.1811 .0552	.0653 .0359	.1502 .0583	.0591 .0201	.0984 .0311	.1922 .1077	

NOTE.—In each cell above the diagonal, three values are shown—χ^2 (top row), df (middle row), and *P*; in each cell below the diagonal, two values are shown—genetic distance (top row) and SE (bottom row).

[a] Difference is significant.

tiveness—e.g., of the *Ava*II RFLP further investigated to the *Hin*cII RFLP—was only .0140 in the Swiss and .0189 in the German sample, since there existed a significant linkage disequilibrium between *Ava*II and *Hin*cII, at least in the Swiss sample.

Linkage Disequilibrium as an Estimate of Physical Distance

$|\Delta|$ Values were significantly correlated with the corresponding physical distances *d* for the Swiss sample ($R = -.358$; $P < .05$), the Italian sample ($R = -.813$; $P < .05$), and the American sample ($R = -.415$; $P < .01$). However, as not expected, relationships between $|\Delta|$ and *d* were not significant for the Germans and Japanese.

Estimations of Recombination Frequencies

A rough estimate of the recombination rate per kilobase, *k*, was calculated from the $|\Delta|$ values by using Mantel's matrix-correspondence method. The overall $4N_ek$ values were 4.107 for the Swiss, 1.465×10^3 for the Germans, 2.963×10^{-1} for the Italians, 1.642×10^1 for the Americans, and 2.867×10^1 for the Japanese. When the effective population size N_e was assumed to be 5,000, then overall *k* values for the different populations were 2.053×10^{-4} for the Swiss, 7.327×10^{-2} for the Germans, 1.482×10^{-5} for the Italians, 8.211×10^{-4} for the Americans, and 1.434×10^{-3} for the Japanese. In figure 2, the estimated regional recombination rates (given as $4N_ek$ values) were higher both between intron 4 (*Taq*I) and exon 8 (*Stu*I) for the Swiss and Germans and between intron 6 (*Sph*I) and exon 8 (*Stu*I) for the Americans. In the Germans and Japanese, the rate was excessively high between exon 13 (*Ava*II) and intron 15 (*Apa*LI [5′]). In the Americans and Swiss, there were increased rates at intron 15, between *Spe*I and *Apa*LI (5′) and between *Apa*LI (5′) and *Pvu*II, respectively. In the Swiss, there was a smaller increase

Table 7

Pairwise Linkage Disequilibrium Values in the Swiss and German Population Samples

	*Taq*I	*Stu*I	*Hin*cII	*Ava*II	*Apa*LI (5′)	*Pvu*II	*Nco*I	*Apa*LI (3′)
*Taq*I		.0702	.1752[a]	−.2039[a]	.2134[a]	−.1235[a]	.0722	.0485
		.3946	.0074	.0021	.0013	.0425	.1272	.1945
		1.0000	.5785	.6804	.6912	.5435	.3378	.2864
*Stu*I	.0815		−.1250	.1232	.0748	−.1990[a]	−.1243	−.0982
	.2335		.0634	.0666	.1632	.0087	.0727	.1772
	.3402		.7101	.7066	.3288	.6436	1.0000	1.0000
*Hin*cII	−.8305[a,b]	−.1803		−.9500[a,b]	−.5347[a,b]	.2662[a,b]	−.1809[a]	−.2724[a,b]
	.0001	.0737		.0001	.0001	.0001	.0044	.0001
	.8939	.7013		.9596	.6910	.4682	.3378	.6436
*Ava*II	−.1917	.0902	.1136		.5067[a,b]	−.2601[a,b]	.1529[a]	.2428[a,b]
	.0775	.5171	.2169		.0001	.0001	.0117	.0003
	1.0000	1.0000	.5520		.6614	.4620	.2878	.5791
*Apa*LI (5′)	.7486[a,b]	−.0815	−.7813[a,b]	−.0021		−.2607[a,b]	.3605[a,b]	.4970[a,b]
	.0001	.2335	.0001	.3407		.0001	.0001	.0001
	.7485	.3402	.8411	.0132		.3549	.5209	.9082
*Pvu*II	−.2784[a]	−.3346[a,b]	.3698[a,b]	−.0857	−.3807[a,b]		.3058[a,b]	.2086[a]
	.0056	.0034	.0003	.2867	.0002		.0001	.0012
	.5764	.6753	.7104	.4717	.7887		.7613	.6573
*Nco*I	.2971[a]	−.0486	−.2535[a]	.0390	.2002[a]	.4105[a,b]		.6786[a,b]
	.0034	.2579	.0102	.3175	.0312	.0001		.0001
	.5055	.1760	.4007	.1758	.3412	.7345		.8587

NOTE.—In each cell, three values are shown: linkage disequilibrium Δ (top row), Fisher's exact probability (middle row), and relative degree of disequilibrium *D′*; data above the diagonal are for eight RFLPs from the Swiss, and data below the diagonal are for seven RFLPs from the German controls.

[a] Significant at $\alpha \leq .05$, by Fisher's exact test.

[b] Significant at the approximation (from Leitersdorf et al. 1989*a*) at $\alpha_i \leq .0018$ (for the Swiss) and at $\alpha_i \leq .0024$ (for the Germans).

between *Stu*I and *Hin*cII; and, in the Germans and Americans, there was a smaller increase between *Hin*cII and *Ava*II. The Italian data revealed no regions with increased recombination frequencies.

Discussion

In the present study, eight RFLPs were selected to cover a large part of the LDL receptor gene, to be informative for segregation analysis, and to be amenable to PCR-based analysis. Thus, the RFLPs for *Taq*I, *Stu*I, *Hin*cII, *Ava*II, and *Nco*I, located in exons or short intron sequences, were first studied using new PCR-based techniques.

To assess relationships between the different populations, three different approaches were used. First, frequencies of restriction sites were compared for each individual RFLP (table 3). Second, frequencies of haplotypes constructed from the common RFLPs of two population samples were directly compared. Third, genetic distances between population samples were calculated. Analysis of frequencies of restriction sites demonstrated a significant heterogeneity among the different population samples, for six of the eight RFLPs (table 3). Direct pairwise comparisons between populations, with respect to their haplotype distributions, revealed differences in 14 of 21 tests (table 6). A certain proportion of significant differences could be due to chance alone, and these results should therefore be interpreted with caution. On the other hand, a lack of significance may be due to a small number of comparable RFLPs—and, thus, of haplotypes; this may have been the case in the comparison of Icelandic haplotypes with those of other Caucasian populations, which yielded no significant differences. Nevertheless, the haplotypes of Icelanders differed significantly from those of the Germans and the Japanese, indicating that

Table 8

Frequency of Informative Cases When Closely Linked RFLPs Are Used for Diagnosing Familial Hypercholesterolemia in Individuals from Switzerland and Germany

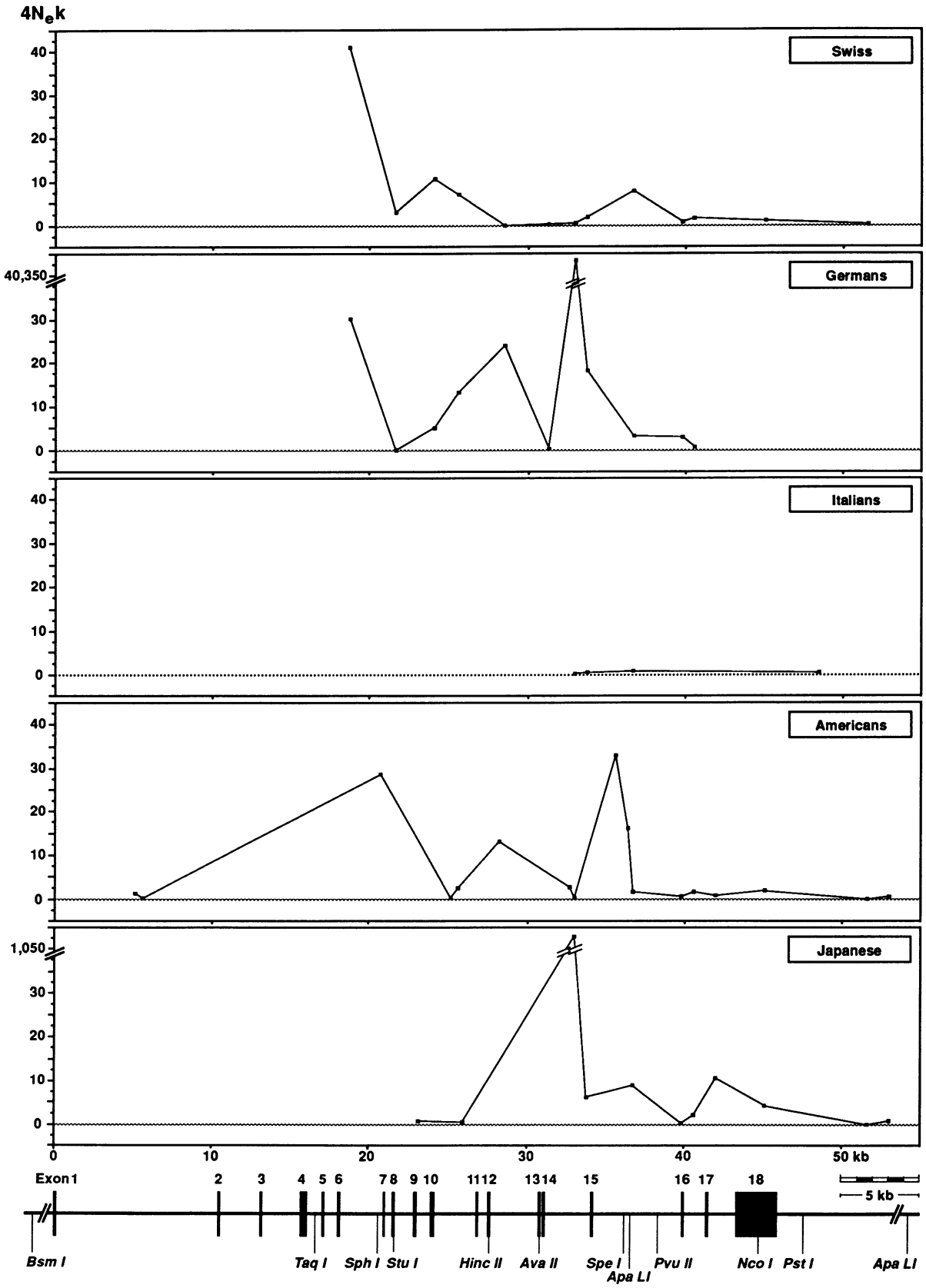
Order of Site Selection	$I_{A\ Swiss}$			$I_{A\ Germans}$	
	RFLP	Individual	Cumulative	RFLP	Individual	Cumulative
1	*Apa*LI (5′)	.3767	.3767	*Apa*LI (5′)	.3832	.3832
2	*Ava*II	.3563	.6187	*Nco*I	.3419	.6504
3	*Pvu*II	.3158	.7683	*Hin*cII	.3794	.7779
4	*Nco*I	.3101	.8257	*Taq*I	.3832	.8204
5	*Taq*I	.2202	.8545	*Pvu*II	.3166	.8413
6	*Apa*LI (3′)	.2721	.8670	*Stu*I	.0919	.8508
7	*Stu*I	.0889	.8774	*Ava*II	.0956	.8588
8	*Hin*cII	.3574	.8812			

in these cases even a small number of comparable RFLPs did not conceal a significant difference. With regard to these compared RFLPs, it is important to note that they were detected in both racial groups (Caucasians and Japanese). Thus, if the RFLPs had not occurred independently more than once, they could be assumed to have arisen before the divergence of the caucasoids and the mongoloids, approximately 40,000 years ago (Nei and Roychoudhury 1982).

At the LDL receptor locus, genetic distances (table 6) between Caucasian and Japanese samples were larger than those among Caucasians, as was expected from investigations of other genetic loci (Cavalli-Sforza and Bodmer 1971, pp. 704–708; Nei and Roychoudhury 1982). In addition, all direct haplotype comparisons demonstrated highly significant differences between the Caucasian and the Japanese samples. However, an exception was the German sample, with larger genetic distances to other Caucasians than to the Japanese. Direct haplotype comparisons between the Germans and the other Caucasians, which showed highly significant differences in all cases, supported this conclusion. Genetic distances were particularly small between North American Caucasians and the Caucasians from Switzerland, Italy, or South Africa. These findings were confirmed by the nonsignificant results of their direct haplotype comparisons.

However, the present results obtained from population samples should be interpreted with caution. First, some of the populations studied in this paper are not geographically or culturally isolated (as are, e.g., Icelanders or Afrikaners; Taylor et al. 1989), and thus the number of investigated individuals may be too small to reflect genetic heterogeneity within a population. In addition, the selection of the normocholesterolemic control subjects was not random, since it was focused on particular centers (e.g., areas of Basel and Munich). Thus, the subjects studied may not be representative of the entire population, and the results may reflect regional differences rather than differences between populations of particular countries. In other respects, the selection of controls was biased only by the fact that they were married to hypercholesterolemic members of FH families; hence, from this point of view, they may be representative of the general populations. In order to reduce the first problem, only families were included which appeared to be autochthonous. As shown in tables 1 and 2, the vast majority (94.1% in the Swiss and 82.0% in the Germans) of independent control haplotypes were observed several times in the samples, and thus they may be more characteristic of the respective populations. Furthermore, in each of these two population samples, one particular haplotype contributed approximately 25% of all control haplotypes. Hence, this

Figure 2 Recombination rates per kilobase, at different regions of the LDL receptor gene, estimated by calculation of $4N_ek$ values from the Δ values of two adjacent RFLPs and, in order to smooth the graph, from the Δ values of two RFLPs with one RFLP in between. The resulting $4N_ek$ values were placed in the middle of the physical distances between the two RFLPs compared, and thus the $4N_ek$ values were aligned with the respective regions of the LDL receptor gene. Lines connecting the $4N_ek$ values were drawn, and the resulting patterns of the regional $4N_ek$ values throughout the LDL receptor genes from five population samples are presented.


4Nek
Swiss
Germans
Italians
Americans
Japanese
40,350
1,050
0
10
20
30
40
50 kb
Exon 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
5 kb
Bsm I
Taq I
Sph I
Stu I
Hinc II
Ava II
Spe I
Apa LI
Pvu II
Nco I
Pst I
Apa LI

haplotype may be typical of that population. The fact that these most frequent haplotypes in the Swiss and Germans differed in three of the seven common RFLPs supports the hypothesis that there may be a real genetic difference with respect to the LDL receptor gene. This is also suggested by the data in table 6, although the geographic distance between Basel and Munich is relatively short (300 km), and the difference may not be sufficiently explained by the divergence, approximately 1,500 years ago, between the ancestors of the German-speaking Swiss (Alemans) and the ancestors of Germans from the Munich area (Bavarians).

To determine how the degree of polymorphism at the marker locus influences the probability of detection of linkage to the index locus, PIC values were calculated. However, since there exists extensive linkage disequilibrium between the different RFLPs, the order in which the sites should be chosen is not only based on the informativeness of a RFLP per se but also on its joint utility with other RFLPs. The observation that the *Apa*LI 5′ RFLP ranked highest in terms of informativeness is in accord with findings by Leitersdorf et al. (1989*a*) in North American Caucasians. The second-ranking RFLP was either *Ava*II or *Hin*cII (Swiss, Germans, and North Americans), and the third-ranking RFLP was either *Nco*I (Germans), *Pvu*II (Swiss), or *Apa*LI 3′ (North Americans). The *Taq*I and *Stu*I RFLPs seemed to be of lower informativeness in these three Caucasian populations without founder effects, studied hitherto. However, in selected pedigrees, such as those in the Japanese and South African Caucasians, these two RFLPs may be crucial for confirmation of FH. In the Swiss sample, the *Ava*II and the *Hin*cII RFLPs showed nonrandom association. In Germans, linkage disequilibrium between *Ava*II and *Hin*cII was not significant. However, in both population samples, the assessment of *Ava*II in addition to *Hin*cII increased the cumulative informativeness only slightly. Hence, in a Swiss or German sample, the simultaneous determination with both RFLPs is, in the majority of cases, not more helpful for segregation analysis than is the determination with one of the two RFLPs.

An approach to study Δ values by computing the corresponding relative linkage disequilibrium D' (table 7) has been described by Leitersdorf et al. (1989*a*). For the two RFLPs with the lowest heterozygosities (table 3)—i.e., *Taq*I and *Stu*I in the Swiss and *Stu*I and *Ava*II in the Germans—values were not significant. However, the relative disequilibrium values were $D' = 1$ in both cases, and thus the calculated Δ's were the maximum values attainable. Two correlations—i.e., (1) of *Stu*I with other RFLPs, in both populations, and (2) of *Ava*II with other RFLPs, in the Germans—revealed no significant disequilibria, even though three of the combinations with *Stu*I in the Swiss and two of the combinations with *AvaII* in Germans attained the maximum D'. In accordance with the findings of Leitersdorf et al. (1989*a*) in North American Caucasians, the present data demonstrated a lack of disequilibrium between RFLPs with low heterozygosities (i.e., *Taq*I, *Stu*I, and *Ava*II) and, hence, low PICs. Thus, extensive linkage disequilibrium throughout large parts of the investigated LDL receptor gene segment is suggested if the Δ values and, in RFLPs with low heterozygosity indexes, the respective relative linkage disequilibria D' are taken into consideration.

Interpretation of linkage disequilibrium Δ and relative disequilibrium D' values has to consider that they are influenced by several genetic processes, including selection, migration, drift, mutation, and recombination. Under uniform recombination, the absolute value of linkage disequilibrium $|\Delta|$ is negatively correlated to the physical distance d for any pair of RFLPs. However, the use of standard linear regression analysis to estimate recombination rates (Chakravarti et al. 1984) may not be strictly applicable to RFLPs, particularly if small gene segments and samples are studied. Problems may arise because the linkage disequilibrium values are not independent, and the approximation may be subject to sampling errors (Weir and Hill 1986). Nevertheless, the formula to assess the relationship between $|\Delta|$ and d has been applied in the investigation of several gene regions, and its use has been supported by direct observations (Chakravarti et al. 1984, 1986; Treco et al. 1985; Chakraborty et al. 1987). Litt and Jorde (1986) proposed Kendall's rank-order correlation test to avoid the problem of a lack of independence of the different $|\Delta|$ values. However, with this method it is still not possible to assess the statistical significance of the resulting correlation coefficients between $|\Delta|$ values and physical distances d. Hence, we used Mantel's matrix-correspondence method to calculate the $4N_e k$ values and correlation coefficients (Chakraborty et al. 1987), and, in addition, we developed a permutational test to evaluate their statistical significance. The expected decrease of linkage disequilibrium with increasing physical distance could be demonstrated in some but not all of the population samples studied in the present paper. A lack of correlation between linkage disequilibrium and physical distance has been observed in at least six

other human genes (β-globin gene [Antonarakis et al. 1982], heavy immunoglobulin chain gene [Bech-Hansen et al. 1983], chromosome 11p [Barker et al. 1984], highly polymorphic region on chromosome 2q [Litt and Jorde 1986], apoAI-apoCIII-apoAIV gene [Antonarakis et al. 1988; Haviland et al. 1991], and apolipoprotein B gene [Zerba et al. 1991]). A first factor which might impair this correlation is genetic drift, particularly if $N_e c$ is small. Second, the missing correlation could be explained on the background of the extreme rarity of recombinations usually observed in small DNA regions. Nonuniform recombination, previously described for the human β-globin gene cluster (Chakravarti et al. 1984), and an excessively high rate of point mutations may be further reasons for the lack of the expected correlation between $|\Delta|$ and d. Finally, estimates of $|\Delta|$ values, calculated for small regions of the genome, may be susceptible to statistical errors (Litt and Jorde 1986). However, the present results of missing correlations in the German and Japanese samples may not be explained either by genetic drift or by a low overall recombination rate: the maximum distance between two RFLPs in the LDL receptor gene region studied in the present paper was approximately 40 kb. Since the human genome is approximately 3×10^6 kb, or 33 Morgan units, in length (Rennwick 1971) and since one recombinational event during meiosis is expected in a DNA fragment of approximately 9.1×10^4 kb, recombination should be observed in a 40-kb interval once in 2,275 meioses, and the expected recombination rate c is approximately 4.4×10^{-4}, since $c = kd$ for small distances. The decay rate of linkage disequilibrium in a population is $D_t = D_0(1 - c)^t$, where D_t is the disequilibrium at generation t and D_0 is the initial linkage disequilibrium at generation 0 (Falconer 1989, pp. 7–21). Thus, t = 1,575 generations, or approximately 40,000 years, would be required to reduce the initial linkage disequilibrium D_0 to half its value for a 40-kb gene segment. Under these conditions, recombinations may not be frequent enough to overcome the effects of new mutations in generating linkage disequilibrium, and the lack of correlation between $|\Delta|$ and d could be explained by a rarity of recombinations in small DNA segments. However, the expected recombination rate ($c \approx 4.4 \times 10^{-4}$) is an underestimate for most of the population samples studied—except for the Italians, where the recombination rate agreed closely with this theoretical value. As deduced from the infinite-sites model (Hudson and Kaplan 1985) and described by Leitersdorf et al. (1989*a*), the 34 different haplotypes in the Swiss and the 24 haplotypes in the German sample are obviously in excess, and more than 20 haplotypes in the Swiss and 10 haplotypes in the Germans might be the result of either an excessively high mutation rate or multiple recombinational events at the LDL receptor locus. In the Swiss, the estimated recombination rate was approximately 19 times higher than theoretically expected, and hence this increased rate may have contributed to the abundant number of haplotypes. Furthermore, in the Germans, the estimated recombination frequency was approximately 357 times larger than that in the Swiss sample and was approximately 6,660 times larger than expected. In the Japanese sample the rate was 652 times larger than expected. Thus, a large number of the haplotypes exceeding the expected ones may have been generated by recombinational events with hot spots of recombination, e.g., between exon 13 and intron 15 (fig. 2). Since clusters of *Alu*-repeats may be the reason for an increased rate of recombination and since the LDL receptor gene contains several *Alu*-repetitive DNA sequences (Yamamoto et al. 1984), e.g., at intron 14 (Horsthemke et al. 1987), the hypothesis that *Alu*-repeats are recombinational hot spots at the LDL receptor locus (Lehrman et al. 1985, 1987*a*; Hobbs et al. 1986) is supported at least by the data from the German and Japanese samples. In the light of these results, it is suggested that not a low recombination rate but, rather, the heterogeneity of recombination frequencies throughout the gene has contributed to the lack of significant correlations between $|\Delta|$ and d values in two of the population samples.

Since, in analyzing the linkage phase in our FH families, we assumed that there were no recombinations, the presence of a high recombination rate might, theoretically, impair the use of the RFLP method in confirming FH. However, despite a high recombination rate, the probability of errors due to recombinational events within the alleles used to follow inheritance through families may be low. This is explained first by the fact that at least seven RFLPs covering a large part of the LDL receptor gene were determined, in order to analyze segregation, and, second, because cosegregation was confirmed in most of the German families by LDL binding studies. Since only families were included in which the resolution of the linkage phase—and in which confirmation of FH thus was possible—families with recombinational events may have been excluded exactly for this reason. If a recombination rate of $k \approx 2.053 \times 10^{-4}$/kb in the Swiss sample is assumed, then one recombination in 122 meioses has to be ex-

pected in a 40-kb DNA segment. Hence, in the 61 Swiss families, comprising 442 meioses, 3.6 recombinations may have occurred. Thus, the result of three families (in which FDB was excluded) demonstrating noncosegregation between FH phenotypes and a particular haplotype may be explained by recombinational events within these families. Thereby, when eight RFLPs are used, recognition of noncosegregation within a given family seems not to be problem, in most cases. However, if only a small number of RFLPs are studied, then recombinations may be a source of serious errors.

As discussed above, the lack of significant correlations between linkage disequilibria $|\Delta|$ and physical distances d, as detected in this study, may be explained by multiple factors. For small regions of the human genome, nonuniform recombination or multiple mutations could generate substantial linkage disequilibrium and may outweigh the effects of uniform recombination. Thus, linkage disequilibrium is probably not always a reliable estimate of the gene order, and caution should be exercised in interpreting $|\Delta|$ values from small gene segments, particularly if only few data are available. However, discrepancy of significant linkage disequilibria, on one hand, and excessively high overall recombination rates usually reducing linkage disequilibrium between RFLPs, on the other hand, are not explained. The present study demonstrated a distinct variety of overall recombination rates among the investigated population samples. Furthermore, heterogeneous recombination rates were revealed throughout the LDL receptor locus, with a different pattern for each population and with hot spots of recombination. Hence, nonuniform recombination may be a mechanism explaining both (1) the lack of correlation between $|\Delta|$ and d and (2) the high average recombination rates. Moreover, recombinational and mutational hot spots are expected to explain both the excess of haplotypes in particular populations and the more than 100 mutations causing FH. Thus, the present study suggests heterogeneity of haplotypes and recombination rates at the LDL receptor locus among different populations and provides new insights into the evolution of the large number of different haplotypes at this clinically crucial gene region.

Acknowledgments

We are grateful to Prof. D. Pometta, Dr. R. James (Geneva), and Dr. R. Darioli (Lausanne) for providing blood samples of patients; to Dr. R. Bannwart (Basel) for help in synthesizing nucleotide primers; to Dr. A. W. A. Hahn (Basel) and Dr. M. Gebhard (Bern) for comments and advice; and to Dr. P. Jordan (Basel) for statistical support. This work was supported by grants from the Postgraduate Course for Experimental Biology and Medicine of the University of Zurich; by grants from Merck, Sharp & Dohme-Chibret (Glattbrugg); and by the "Wissenschaftlicher Kredit" of the University Hospital of Basel.

References

Aalto-Setälä K, Helve E, Kovanen PT, Kontula K (1989) Finnish type of low density lipoprotein receptor gene mutation (FH-Helsinki) deletes exons encoding the carboxy-terminal part of the receptor and creates an internalization-defective phenotype. J Clin Invest 84:499–505

Antonarakis SE, Boehm CD, Giardina PJV, Kazazian HH (1982) Nonrandom association of polymorphic restriction sites in the β-globin gene cluster. Proc Natl Acad Sci USA 79:137–141

Antonarakis SE, Oettgen P, Chakravarti A, Halloran SL, Hudson RR, Feisee L, Karathanasis SK (1988) DNA polymorphism haplotypes of the human apolipoprotein apoA1-apoC3-apoA4 gene cluster. Hum Genet 80:265–273

Barker D, Holm T, White R (1984) A locus on chromosome 11p with multiple restriction site polymorphisms. Am J Hum Genet 36:1159–1171

Bech-Hansen NT, Linsley PS, Cox DW (1983) Restriction fragment length polymorphisms associated with immunoglobulin C_γ genes reveal linkage disequilibrium and genomic organization. Proc Natl Acad Sci USA 80:6952–6956

Bland M (1989) An introduction in medical statistics. Oxford University Press, Oxford, pp 245–247

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Brink PA, Steyn LT, Coetzee GA, Van der Westhuyzen DR (1987) Familial hypercholesterolemia in South African Afrikaners: Pvu II and Stu I DNA polymorphisms in the LDL-receptor gene consistent with a predominating founder gene effect. Hum Genet 77:32–35

Brown MS, Goldstein JL (1991) The hyperlipoproteinemias and other disorders of lipid metabolism. In: Wilson JD, Braunwald E, Isselbacher KJ, Petersdorf RG, Martin JB, Fauci AS, Root RK (eds) Principles of internal medicine, 12th ed. McGraw-Hill, New York, pp 1814–1825

Cavalli-Sforza LL, Bodmer WF (1971) The genetics of human populations. WH Freeman, San Francisco

Chakraborty R, Lidsky AS, Daiger SP, Güttler F, Sullivan S, Dilella AG, Woo SLC (1987) Polymorphic DNA haplotypes at the human phenylalanine hydroxylase locus and their relationship with phenylketonuria. Hum Genet 76:40–46

Chakravarti A, Buetow KH (1985) A strategy for using multiple linked markers for genetic counseling. Am J Hum Genet 37:984–997

Chakravarti A, Buetow KH, Antonarakis SE, Waber PG, Boehm CD, Kazazian HH (1984) Nonuniform recombination within the human β-globin gene cluster. Am J Hum Genet 36:1239–1258

Chakravarti A, Elbein SC, Permutt MA (1986) Evidence for increased recombination near the human insulin gene: implication for disease association studies. Proc Natl Acad Sci USA 83:1045–1049

Daga A, Mattioni T, Balestreri R, Coviello DA, Corte G, Bertolini S (1990) Use of three DNA polymorphisms of the LDL receptor gene in the diagnosis of familial hypercholesterolemia. Hum Genet 84:412–416

Falconer DS (ed) (1989) Introduction to quantitative genetics, 3d ed. Longman, Harlow, UK

Haviland MB, Kessling AM, Davignon J, Sing CF (1991) Estimation of Hardy-Weinberg and pairwise disequilibrium in the apolipoprotein AI-CIII-AIV gene cluster. Am J Hum Genet 49:350–365

Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231

Hobbs HH, Brown MS, Goldstein JL, Russel DW (1986) Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia. J Biol Chem 261:13114–13120

Hobbs HH, Brown MS, Russel DW, Davignon J, Goldstein JL (1987*a*) Deletion in the gene for the low-density-lipoprotein receptor in a majority of French Canadians with familial hypercholesterolemia. N Engl J Med 317:734–737

Hobbs HH, Esser V, Russel DW (1987*b*) Ava II polymorphism in the human LDL receptor gene. Nucleic Acids Res 15:379

Hobbs HH, Leitersdorf E, Goldstein JL, Brown MS, Russell DW (1988) Multiple crm$^-$ mutations in familial hypercholesterolemia: evidence for 13 alleles, including four deletions. J Clin Invest 81:909–917

Horsthemke B, Beisiegel U, Dunning A, Havinga JR, Williamson R, Humphries S (1987) Unequal crossing over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. Eur J Biochem 164:77–81

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164

Humphries SE, Kessling AM, Horsthemke B, Donald JA, Seed M, Jowett N, Holm M, et al (1985) A common DNA polymorphism of the low-density lipoprotein (LDL) receptor gene and its use in diagnosis. Lancet 1:1003–1005

Humphries SE, Taylor R, Jeenah M, Dunning A, Horsthemke B, Seed M, Schuster H, et al (1989) Gene probes in diagnosis of familial hypercholesterolemia. Arteriosclerosis 9, Suppl I:I-59–I-65

Ingram D, Bloch RF (eds) (1984) Mathematical methods in medicine. I. Statistical and analytical techniques, 1st ed. John Wiley & Sons, Chichester

Kotze MJ, Langenhoven E, Dietzsch E, Retief AE (1987) A RFLP associated with the low-density lipoprotein receptor gene (LDLR). Nucleic Acids Res 15:376

Kotze MJ, Langenhoven E, Retief AE, Seftel HC, Henderson HE, Weich HFH (1989*a*) Haplotypes identified by 10 DNA restriction fragment length polymorphisms at the human low density lipoprotein receptor gene locus. J Med Genet 26:255–259

Kotze MJ, Langenhoven E, Warnich L, du Plessis L, Marx MP, Oosthuizen CJJ, Retief AE (1989*b*) The identification of two low-density lipoprotein receptor gene mutations in South African familial hypercholesterolemia. S Afr Med J 76:399–401

Kotze MJ, Langenhoven E, Warnich L, Marx MP, Retief AE (1989*c*) Molecular characterisation of a low-frequency mutation in exon 8 of the human low-density lipoprotein receptor gene. S Afr Med J 76:402–405

Kotze MJ, Retief AE, Brink PA, Weich HFH (1986) A DNA polymorphism in the human low-density lipoprotein receptor gene. S Afr Med J 70:77–79

Lehrman MA, Goldstein JL, Russel DW, Brown MS (1987*a*) Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. Cell 48:827–835

Lehrman MA, Schneider WJ, Brown MS, Davis CG, Elhammer A, Russel DW, Goldstein JL (1987*b*) The Lebanese allele at the low density lipoprotein receptor locus: nonsense mutation produces truncated receptor that is retained in endoplasmic reticulum. J Biol Chem 262:401–410

Lehrman MA, Schneider WJ, Südhof TC, Brown MS, Goldstein JL, Russel DW (1985) Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. Science 227:140–146

Leitersdorf E, Chakravarti A, Hobbs HH (1989*a*) Polymorphic DNA haplotypes at the LDL receptor locus. Am J Hum Genet 44:409–421

Leitersdorf E, Hobbs HH (1987) Human LDL receptor gene: two ApaL I RFLPs. Nucleic Acids Res 15:2782

——— (1988) Human LDL receptor gene: Hinc II polymorphism detected by gene amplification. Nucleic Acids Res 16:7215

Leitersdorf E, Van Der Westhuyzen DR, Coetzee GA, Hobbs HH (1989*b*) Two common low density lipoprotein receptor gene mutations cause familial hypercholesterolemia in Afrikaners. J Clin Invest 84:954–961

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49:49–67

Litt M, Jorde LB (1986) Linkage disequilibria between pairs of loci within a highly polymorphic region of chromosome 2Q. Am J Hum Genet 39:166–178

Ma YH, Bétard C, Roy M, Davignon J, Kessling AM (1989) Identification of a second "French Canadian" LDL receptor gene deletion and development of a rapid method to detect both deletions. Clin Genet 36:219–228

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

Nei M, Roychoudhury AK (1982) Genetic relationship and evolution of human races. In: Hecht MK, Wallace B, Prance GT (eds) Evolutionary biology. Plenum, New York, pp 1–59

Ohta T, Kimura M (1969) Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics 63:229–238

Rennwick JH (1971) The mapping of human chromosomes. Annu Rev Genet 5:81–120

Schuster H, Rauh G, Müller S, Keller C, Wolfram G, Zöllner N (1992) Allele-specific and asymmetric polymerase chain reaction amplification in combination: a one step polymerase chain reaction protocol for rapid diagnosis of familial defective apolipoprotein B-100. Anal Biochem 204:22–25

Schuster H, Stiefenhofer B, Wolfram G, Keller C, Humphries S, Huber A, Zöllner N (1989) Four DNA polymorphisms in the LDL-receptor gene and their use in diagnosis of familial hypercholesterolemia. Hum Genet 82:69–72

Soria LF, Ludwig EH, Clarke HRG, Vega GL, Grundy SM, McCarthy BJ (1989) Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. Proc Natl Acad Sci USA 86:587–591

Spengel FA, Harders-Spengel KM, Keller CF, Wieczorek A, Wolfram G, Zöllner N (1982) Use of fibroblast culture to diagnose and genotype familial hypercholesterolaemia. Ann Nutr Metab 26:240–247

Südhof TC, Goldstein JL, Brown MS, Russell DW (1985) The LDL receptor gene: a mosaic of exons shared with different proteins. Science 228:815–822

Sved JA (1971) Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor Popul Biol 2:125–141

Taylor R, Bryant J, Gudnason V, Sigurdsson G, Humphries S (1989) A study of familial hypercholesterolemia in Iceland using RFLPs. J Med Genet 26:494–498

Treco D, Thomas B, Arnheim N (1985) Recombination hot spot in the human β-globin gene cluster: meiotic recombination of human DNA fragments in *S. cerevisiae*. Mol Cell Biol 5:2029–2030

Weir BS, Hill WG (1986) Nonuniform recombination within the human β-globin gene cluster. Am J Hum Genet 38:776–778

Yamakawa K, Okafuji T, Iwamura Y, Russel DW, Hamaguchi H (1987) Taq I polymorphism in the human LDL receptor gene. Nucleic Acids Res 15:7659

Yamakawa K, Yanagi H, Saku K, Sasaki J, Okafuji T, Shimakura Y, Kawai K, et al (1991) Family studies of the LDL receptor gene of relatively severe hereditary hypercholesterolemia associated with Achilles tendon xanthomas. Hum Genet 86:445–449

Yamamoto T, Davis CG, Brown MS, Schneider WJ, Casey ML, Goldstein JL, Russel DW (1984) The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. Cell 39:27–38

Zerba KE, Kessling AM, Davignon J, Sing CF (1991) Genetic structure and the search for genotype-phenotype relationships: an example from disequilibrium in the Apo B gene region. Genetics 129:525–533