# Haplotypes identified by <sup>10</sup> DNA restriction fragment length polymorphisms at the human low density lipoprotein receptor gene locus

M <sup>J</sup> KOTZE\*, E LANGENHOVEN\*, A E RETIEF\*, H C SEFTELt, H E HENDERSON<sup>‡</sup>, AND H F H WEICH§

From \*the Research Unit for Cytogenetics, Department of Human Genetics, University of Stellenbosch Medical School, Tygerberg; tthe Department of Medicine, University of Witwatersrand, Johannesburg; #he Department of Chemical Pathology, Red Cross War Memorial Children's Hospital, Cape Town; and §the Department of Internal Medicine and Bayer Cardiovascular Clinical Research Unit, Tygerberg Hospital, Tygerberg 7505, South Africa.

SUMMARY Ten useful two allele restriction fragment length polymorphisms of the low density lipoprotein receptor gene were used for haplotype analysis in 45 unrelated familial hypercholesterolaemic (FH) patients, <sup>60</sup> normal controls, and <sup>32</sup> FH homozygotes, all of whom were white Afrikaners. Pedigree analysis in <sup>27</sup> informative heterozygous FH and <sup>23</sup> normal families has shown the segregation of at least 17 haplotypes in the normal population (111 chromosomes) compared to <sup>a</sup> predominant association of two of these haplotypes with the disease in the FH subjects. This association was further confirmed in <sup>32</sup> FH homozygotes, indicating at least two 'founder' members for the disease in the Afrikaner population. Recombination events were not detected in any of the families studied and we thus conclude that the haplotypes associated with FH function as specific markers for the disease and will allow presymptomatic diagnosis in affected families.

The prevalence of familial hypercholesterolaemia (FH), an autosomal dominant disease characterised by raised low density lipoprotein (LDL) cholesterol levels, is five to 10 times higher in the Afrikaner population than has been reported for most other population groups in the world.<sup>1 2</sup> Seftel *et al*<sup>1</sup> have suggested a 'founder gene effect' to explain this abnormally high frequency. Clinical studies have indicated that tendon xanthomas and premature coronary artery disease (CAD) are encountered in middle aged FH heterozygotes, while in the homozygous condition both cutaneous and tendon xanthomas as well as aortic and coronary atherosclerosis generally appear before the age of  $10$  years.<sup>3</sup>

Over recent years it has been shown by Brown and Goldstein<sup>4</sup> that FH is caused by mutations in the LDL receptor gene. The availability of cDNA and genomic clones of this gene have now made it possible to characterise mutations at the DNA level.4 <sup>5</sup> We have previously provided evidence for the existence of at least two 'founder' FH genes in

Received for publication 26 September 1988. Accepted for publication 4 October 1988.

# Subjects and methods

The FH heterozygous patients and normocholesterolaemic subjects were those previously described.6 Pedigree analyses were performed on 27 of the 45 heterozygous FH patients and 29 of the <sup>60</sup> normal subjects. Blood samples were collected from 292 of the FH family members and <sup>130</sup> of the members from normal families. Thirty-two patients with homozygous familial hypercholesterolaemia from 27

the Afrikaner population.<sup>6</sup> Four haplotypes defined by the restriction enzymes  $PvuII$  (P), Stul (S), and NcoI (N) were found to segregate in this population. In 71% of FH families studied, the defective gene cosegregated with the rare allele of a NcoI RFLP<sup>7</sup> (P1S1N2 haplotype), while the rare allele of a  $StuI$ RFLP8 (P2S2N1 haplotype) segregated with FH in 20% of these families. We now report on haplotype associations of <sup>10</sup> different DNA polymorphisms at the LDL receptor gene locus in the normocholesterolaemic and FH populations and in <sup>a</sup> large number of FH homozygotes.

white Afrikaner families, attending lipid clinics at the Tygerberg, Johannesburg, Groote Schuur, and Red Cross War Memorial Children's Hospitals, South Africa, donated blood samples for DNA analysis.

The homozygous state of FH was diagnosed on the basis of the following criteria<sup>1</sup>:  $(1)$  plasma cholesterol concentrations greater than  $15 \text{ mmol/}$ : (2) the appearance of xanthomas in the first decade of life; and (3) hypercholesterolaemia or early CAD in the parents of the proband.

Methods used in determination of lipoproteins, preparation of DNA, restriction endonuclease digestion, transfer to nitrocellulose, and prehybridisation to nick translated DNA probes have been described elsewhere.<sup>6</sup> Filters were washed as described with a final wash in  $1 \times SSC$  (0.15 mol/l NaCl, 0.015 mol/l Na citrate) at 65°C. Three human DNA probes for the LDL receptor gene were used in this study. Probe 1, pLDLR-2HH1, consists of a  $1.9$  kb fragment of the  $3'$  end of the LDL receptor cDNA clone9 and was obtained from D W Russell of Dallas. Probe 2 was a 1-05 kb PstI fragment

subcloned from pLDLR-3, <sup>a</sup> full length LDL receptor cDNA; pLDLR-3 was obtained from the American type culture collection (ATCC). Probe 3 was a 3.8 kb ApaI fragment of pLDLR-3. The RFLPs determined were those for the restriction enzymes TaqI,  $^{10}$  StuI,  $^{8}$  BstEII,  $^{11}$   $^{12}$  AvaII,  $^{13}$  ApaLI,  $^{14}$ PvuII,<sup>15</sup> NcoI,' MspI,<sup>16</sup> and PstI.<sup>17</sup>

 $\gamma^2$  tests were performed to determine whether the populations investigated were in Hardy-Weinberg equilibrium and to compare the allele frequencies.

## Results

#### CLINICAL DATA

Results on lipid analysis, family history of early CAD, and presence of tendon xanthomas in the FH heterozygotes and control subjects have been described elsewhere.6 Table <sup>1</sup> summarises the clinical data on the FH homozygotes.

## ALLELE FREQUENCIES IN FH AND NORMOCHOLESTEROLAEMIC POPULATIONS The frequencies of the LDL receptor alleles charac-

<b>Kindred</b>	Patient No	Age (y)	Sex	Cholesterol (mmol/l)			<b>Total</b> triglycerides (mmol/l)	XMTA*	<b>CADt</b>
				Total	LDL	<b>HDL</b>			
1	8	32	F	16.5	$15-4$	0.9	$1 - 1$	$\ddot{}$	$\ddot{}$
$\mathbf{2}$	35	28	F	$19-2$	$17-6$	$1-3$	1.3	$\ddot{}$	
3	109	14	F	$15 - 7$	$15 - 3$	$0 - 1$	$1-3$	$\ddot{}$	
4	225	13	F	$17-6$	$16-5$	$1-0$	0.7	+	
	226	7	F	$20 - 4$	19.3	0.9	$1 - 1$	+	
5	253	7	F	$21 - 8$	$20 - 6$	0.9	$1 - 4$	$\ddot{}$	
6	254	3	F	25.5	$24 - 4$	$0 - 8$	1.5	$\ddot{}$	
7	305	11	M	$22 - 7$	$21 - 7$	$0 - 8$	1.3	+	
8	383	4	F	$21 - 1$	$19 - 8$	$1 - 1$	$1-0$	+	
9	387		F	$17 - 7$	$16 - 1$	$1-2$	1.9	$\ddot{}$	
10	401	6	F	$15 - 2$	$14 - 4$	$0 - 6$	0.9	+	$\ddot{}$
11	610	$\overline{c}$	F	$15 - 7$	$14-6$	0.8	$1-6$	$\ddot{}$	
12	834	19	M	$17 - 5$	$16 - 6$	0.7	0.5	+	
	835	16	M	23.9	$23 - 1$	0.6	$1 - 4$	$\ddot{}$	
13	836	12	F	$19-6$	$18 - 1$	0.9	$1-4$	+	
14	837	40	M	$19 - 2$	17.9	$0 - 6$	3.2	$\ddot{}$	$\pmb{+}$
15	842	27	M	$16 - 5$	$16 - 0$	$0 - 4$	0.3	+	+
16	863	19	F	$17 - 8$	$16-9$	0.7	0.4	+	
17	866	24	M	$22 - 6$	$21 - 7$	0.6	$1 - 4$	$\ddot{}$	
	867	19	F	$24 - 6$	$23 - 7$	$0 - 6$	1.6	$\ddot{}$	
18	869	19	F	$16 - 7$	$15 - 8$	$0 - 4$	$1-0$	$\ddot{}$	$\ddot{}$
	872	16	M	$18 - 4$	$17 - 4$	0.5	0.9	+	
19	875	10	F	$18 - 8$	$18 - 1$	0.4	$0 - 6$	+	
20	878	29	F	$21 - 6$	$20 - 8$	$0 - 6$	0.5	$\ddot{}$	$\ddot{}$
	885	15	F	$19-6$	$18 - 5$	0.9	0.7	$\ddot{}$	
21	840	54	F	$19-6$	$18 - 7$	0.7	$1-4$	+	+
	879	52	F	$20-3$	19.3	0.5	$2 - 1$	$\ddot{}$	$\ddot{}$
22	881	17	F	$18 - 1$	$17-4$	0.5	$1-0$	+	
23	884	14	M	18.9	$17-6$	$1-0$	0.7	$\ddot{}$	
24	887	26	F	$16-9$	$16-1$	0.9	0.6	+	$\ddot{}$
25	890	11	M	$19-7$	18.5	0.7	$1-0$	+	
26	898	10	M	$16 - 6$	$15-4$	0.5	1.3	+	
27	904	1	F	$25 - 7$	$24 - 8$	0.5	2.0	$\ddot{}$	

TABLE <sup>1</sup> Clinical and biochemical features of homozygous FH patients.

\*Planar cutaneous xanthomas.

tCoronary artery disease.

terised by the 10 different polymorphic sites have been calculated for 60 normocholesterolaemic and <sup>45</sup> unrelated FH heterozygotes (table 2). All variants showed autosomal codominant inheritance and the distribution of homozygotes and heterozygotes at each site was consistent with Hardy-Weinberg equilibria determined from the individual allele frequencies. We previously reported <sup>a</sup> statistically significant excess of the rare allele of the NcoI RFLP in our FH population  $(\chi^2=17.2, 1 \text{ df}, p<0.005)$ . An increase has also been found in the rare allele of the ApaLI (3') RFLP ( $\chi^2$ =59·1, p<0·005), and common alleles of the Avall  $(\chi^2=6.7, p<0.01)$ , ApaLI (5')  $(\chi^2=14.6, p<0.005)$ , *Pst*1 ( $\chi^2=10.4, p<0.005$ ), and *TaqI* ( $\chi^2=4.2, p<0.05$ ) RFLPs in FH patients compared to normocholesterolaemic controls. This indicates population association of these alleles with the gene predisposing to hypercholesterolaemia.

## COMPARISON OF ALLELE FREQUENCIES AMONG DIFFERENT NORMAL POPULATION GROUPS

We compared the allele frequencies of the <sup>10</sup> RFLPs at the LDL receptor gene in the normal South African population with the frequencies published for populations in other countries. The frequencies of the common and rare alleles of the TaqI RFLP reported for the Japanese<sup>10</sup> differ significantly from those observed in our population  $(\chi^2=33.7, p<0.005)$ (table 2). Differences were also found with alleles defined by the AvaII<sup>13</sup> ( $\chi^2$ =6.5, p<0.025), PvuII<sup>13</sup>  $(\chi^2=4.3, p<0.05)$ , Mspl<sup>10</sup> ( $\chi^2=6.0, p<0.025$ ), and ApaLI  $(3')^{14}$   $(\chi^2=19.7, p<0.005)$  RFLPs reported for Caucasians.

# HAPLOTYPE ANALYSIS IN THE NORMAL POPULATION

By assigning a '1' for the presence of a common RFLP allele and <sup>a</sup> '2' for the rare allele, haplotype arrangements for all 10 polymorphic sites taken as a unit have been determined in 111 normal chromosomes. Haplotype status was assigned by family studies (73 chromosomes) or by homozygosity for at least nine of the 10 RFLPs (38 chromosomes). Haplotypes not associated with the disease in FH subjects and families were also counted as normal. Pedigree information was not available for <sup>10</sup> FH subjects and 22 normal subjects heterozygous for two or more of the RFLPs and were therefore omitted from this analysis.

We observed marked linkage disequilibrium between the <sup>10</sup> polymorphic sites at the LDL receptor gene locus. Since each of these sites distinguishes two alleles, there are  $1024 (2^{10})$  possible haplotypic arrangements. Only 17 of these were detected in the 111 chromosomes studied and are listed in table 3 together with their respective





\*See text.

TABLE 3 Frequencies of haplotypes in 111 chromosomes of the normal population.

<b>Restrictions</b> sites									No of chromosomes	Haplotype No	%	
T	$\pmb{S}$		B A A5' P N M Ps A3'									
			,	2				2		34		30
			1			2	1		2	26		23
					2		2			16		14
		2			2		2			6		5
			$\overline{c}$	$\mathbf{2}$		$\mathbf{2}$	1			n		5
	2	2	1		2		$\overline{c}$			5	n	
		2								3		3
			1			$\overline{c}$				3	8	3
2		2	1	2						2	9	2
							2			2	10	2
						2				,	11	2
	2	2	1				2		2		12	
			2	2				2	2		13	
			2	$\overline{2}$			2				14	
				$\overline{2}$		1	$\overline{c}$				15	
2				2		1	2				16	
2					$\overline{c}$			2			17	

frequencies. Arbitrary numbers have been allocated to the haplotypes found.

### SEGREGATION OF THE RFLP HAPLOTYPES IN FH FAMILIES

Table 4 shows the frequencies of the different haplotypes associated with FH in the families investigated, together with the haplotypes from a further eight FH subjects who did not warrant pedigree analysis for the haplotype determinations as they were homozygous for all 10 RFLPs. Haplotype <sup>2</sup> was found to segregate with FH in <sup>17</sup> families (46 normal and 75 affected offspring) and seven of

TABLE <sup>4</sup> Haplotypes associated with <sup>35</sup> FH heterozygotes.

<b>Restrictions</b> sites								No of chromosomes	Haplotype No	%		
				T S B A A5' P N M P <sub>S</sub> A3'								
					1	$\overline{2}$	- 1	-1	2	24		69
	$\overline{2}$	$\overline{2}$	$\mathbf{1}$	-1	2 <sub>1</sub>		$\overline{2}$					20
		1	$\overline{2}$	$\mathcal{P}$	1		1	2				3
					າ		,					
		,			2	$\blacksquare$	$\mathfrak{p}$					
		1	2 <sub>2</sub>		1		1	$\mathbf{r}$	2		13	

TABLE <sup>5</sup> Haplotypes found in <sup>32</sup> clinical FH homozygotes.



the RFLP homozygotes (69% of the total FH alleles), while haplotype 6 segregated in six informative families (34 normal and 44 affected offspring) and the remaining RFLP homozygote (20%). Four separate haplotypes (1, 3, 4, and 13) were found in association with the FH phenotype in the remaining four informative families (3%).

The haplotypes determined for <sup>32</sup> FH homozygotes (from 27 kindreds) are shown in table 5. Twenty-nine of these subjects, representing 24 separate kindreds, were homozygous for haplotype 2, while subjects 8, 383, and 884 (table 1) showed compound heterozygosity for haplotypes 2 and 6. These results establish a clear predominance in the FH population studied of the two haplotypes among the 17 observed in the normal South African population.

### **Discussion**

A total of <sup>10</sup> useful two allele probe-enzyme systems, spanning a distance of 45 kb, were used for RFLP haplotype studies at the LDL receptor gene locus. We have shown marked linkage disequilibrium between these 10 markers, since only 17 haplotypes among 1024 possible haplotypic combinations were observed in 111 chromosomes from the normal South African population. This finding is in keeping with other RFLP studies at the immunoglobulin (200 kb) and  $\beta$  globin (60 kb)<sup>18</sup> loci where linkage disequilibrium was also shown over relatively short genetic distances. The major reason for this phenomenon is the close proximity of the polymorphic sites and the resulting low frequency of their separation through recombination events. This was evident in our study of several two and three generation families where no recombinants were detected.

We have found significant differences in the frequency of five of the <sup>10</sup> RFLPs at the LDL receptor locus in South African Afrikaners compared with populations in other countries. This most probably also indicates differing haplotype frequencies. Discrepancies in the number and frequency of haplotypes between populations have been ascribed to selective advantage for certain alleles, or genetic drift associated with the effects of migration. Founder effects also play a role and probably underly the distinguishing RFLP frequencies in our Afrikaner population. A founder effect is well documented in this group which has remained isolated through religious belief and cultural bonds.

Pedigree analyses in normal and FH families have shown the presence of a small number of haplotypes in normal chromosomes. Only two of these haplotypes are associated with FH in the homozygous FH patients and we thus conclude that the defective LDL receptor genes associated with these haplotypes (2 and 6) account for the founder effect in the Afrikaner and probably represent two separate gene defects. Only four of the FH families studied  $(11\%)$ have shown association with other haplotypes (1, 3, 4, and 13), which probably correspond to the 10 times lower FH frequency found in other population groups.

The nature of the alteration in the DNA structure of the LDL receptor genes underlying FH in the Afrikaner population has not yet been reported. Our results clearly indicate that haplotypes function as specific markers for FH in affected families. Early detection of carriers is thus possible, allowing for dietary and life style modification to reduce the subsequent risk of early heart disease.

We thank Professor G M B Berger for the inclusion in the study of FH patients attending the lipid clinics at Groote Schuur and Red Cross Children's Hospitals, Sheila Baker and the nursing staff of Genetic Services, Department of Health and Welfare for collecting the blood samples, and Ina Jordaan for technical assistance; This work was supported by the South African Medical Research Council, the University of Stellenbosch, and the Cape Provincial Administration. The work is part of a thesis to be submitted in fulfilment of the requirements for the degree of Doctor of Philosophy, University of Stellenbosch.

#### References

<sup>1</sup> Seftel HC, Baker SG, Sandler MP, et al. A host of hypercholesterolaemic homozygotes in South Africa. Br Med J 1980;28: 633-6.

- <sup>2</sup> Jooste PL, Benade AJS, Rossouw JE. Prevalence of familial hypercholesterolaemia in three rural South African communities. S Afr Med J 1986;69:548-51.
- <sup>3</sup> Goldstein JL, Brown MS. Familial hypercholesterolaemia. In: Stanbury JB, Wyngaarden JB, Fredrickson DS, Goldstein JL, Brown MS, eds. The metabolic basis of inherited disease. 5th ed. New York: McGraw-Hill, 1983:672-712.
- Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. Science 1986;232:34-47.
- <sup>5</sup> Sudhof TC, Goldstein JL, Brown MS, Russel DW. The LDL receptor gene: a mosaic of exons shared with different proteins. Science 1985;228:815-22.
- <sup>6</sup> Kotze MJ, Langenhoven E, Retief AE, et al. Haplotype associations of three DNA polymorphisms at the human low density lipoprotein receptor gene locus in familial hypercholesterolaemia. J Med Genet 1987;24:750-5.
- <sup>7</sup> Kotze MJ, Langenhoven E, Dietzsch E, Retief AE. A RFLP associated with the low-density lipoprotein receptor gene. Nucleic Acids Res 1987;15:376.
- <sup>8</sup> Kotze MJ, Retief AE, Brink PA, Weich HFH. A DNA polymorphism in the human low-density lipoprotein receptor
- gene. S Afr Med J 1986;70:77-9.<br>Yamamoto T, Davis CG, Brown MS, et al. The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. Cell 1984;39:27-38.
- <sup>10</sup> Yamakawa K, Okafuji T, Iwamura Y, et al. Taq <sup>I</sup> polymorphism in the human LDL receptor gene. Nucleic Acids Res 1987;15:7659.
- <sup>11</sup> Steyn LT, Pretorius A, Brink PA, Bester AJ. RFLP for the human LDL receptor gene (LDLR): Bst ElI. Nucleic Acids Res 1987;11:4702.
- <sup>12</sup> Kotze MJ, Langenhoven E, Retief AE. Improved visualization of the Bst ElI RFLP of the human LDL receptor gene by codigestion. Nucleic Acids Res 1987;23:10067.
- <sup>13</sup> Hobbs HH, Esser V, Russell DW. Ava II polymorphism in the human LDL receptor gene. Nucleic Acids Res 1987;1:379.
- <sup>14</sup> Leitersdorf E, Hobbs HH. Human LDL receptor gene: two Apa LI RFLPs. Nucleic Acids Res 1987;6:2782.
- <sup>15</sup> Humphries SE, Horsthemke B, Seed M, et al. A common DNA polymorphism of the low-density lipoprotein (LDL) receptor
- gene and its use in diagnosis. *Lancet* 1985;i:1003–5.<br><sup>16</sup> Geisel J, Weisshaar B, Oette K, *et al.* Double Msp I RFLP in the human LDL receptor gene. Nucleic Acids Res 1987;9:3943.
- Funke H, Klug J, Frossard P, et al. Pst <sup>I</sup> RFLP close to the LDL receptor gene. Nucleic Acids Res 1986;19:7820.
- 18 Antonarakis SE, Boehm CD, Giardina PVJ, Kazazian HH. Non-random association of polymorphic restriction sites in the betaglobin gene cluster. Proc Natl Acad Sci USA 1982;79:  $137 - 41$ .

Correspondence to Professor A E Retief, Department of Human Genetics, University of Stellenbosch, Medical School, PO Box 63, Tygerberg 7505, South Africa.