

## Mutation in the ARH Gene and a Chromosome 13q Locus Influence Cholesterol Levels in a New Form of Digenic-Recessive Familial Hypercholesterolemia

Hussam Al-Kateb, Sylvia Bähring, Katrin Hoffmann, Konstantin Strauch, Andreas Busjahn, Gudrun Nürnberg, Muhidien Jouma, Eckehard K.F. Bautz, Hans A. Dresel, Friedrich C. Luft

**Abstract**—We studied a Syrian family with 3 children who had low-density lipoprotein cholesterol (LDL) concentrations of 13.3, 12.2, and 8.6 mmol/L, respectively. Three other siblings and the parents all had LDL values <4.52 mmol/L, suggesting an autosomal-recessive mode of inheritance. The extended pedigree had 66 additional persons with normal LDL values. A genome-wide scan in the core family with 427 markers showed support for linkage on both chromosomes 1 and 13. Markers on chromosome 1 revealed a 3.07 multipoint LOD score between 1p36.1-p35, an 18-cM interval. Surprisingly, we also found linkage to 13q22-q32, a 14-cM interval, with a 3.08 LOD score. We had identified this locus earlier as containing a gene strongly influencing LDL in another Arab family with autosomal-dominant familial hypercholesterolemia and in normal dizygotic twins. We found evidence for an interaction between these loci. We next genotyped our twin panel and confirmed linkage of the 1p36.1-p35 locus to LDL ( $P < 0.002$ ) in this normal population. Elucidation of ARH, the LDL receptor adaptor protein at chromosome 1p35, caused us to sequence that gene. We first identified the genomic structure of ARH gene and then sequenced the gene in our family. We found an intron 1 acceptor splice-site mutation. This mutation was not found in any other family members, in 31 nonrelated Syrian persons, or in 30 Germans. Our results underscore the importance of ARH on chromosome 1 and the chromosome 13q locus to LDL, not only in families with unusual illnesses, but also to the general population. (*Circ Res.* 2002;90:951-958.)

**Key Words:** cholesterol ■ LDL ■ hyperlipidemia ■ genetics ■ familial hypercholesterolemia

Elevated plasma low-density lipoprotein cholesterol (LDL) concentrations, an important risk factor for atherosclerosis, are influenced by genetic factors. Elucidation of monogenic lipid disturbances has contributed greatly to our understanding of these disorders specifically and to lipid metabolism in general.<sup>1-4</sup> Ciccarese et al<sup>5</sup> described families from Sardinia in whom hypercholesterolemia is inherited in an autosomal-recessive fashion (ARH). They studied 5 families and were successful in mapping an ARH locus to chromosome 15q25-q26, which they termed ARH-1. Eden et al<sup>6</sup> and this study independently show that the autosomal-recessive familial hypercholesterolemia gene maps to chromosome 1p36.1-p35 (ARH). Recently, Garcia et al<sup>7</sup> succeeded in cloning the gene residing at the chromosome 1p35 locus. The gene codes for a putative LDL receptor adaptor protein. In an earlier study, we found linkage of unexpectedly low LDL values in a family with autosomal-dominant familial hypercholesterolemia (FH) to a locus on chromosome 13q.<sup>8</sup> We now report novel findings in an Arab kindred with ARH. We describe linkage to both the chromosome 1p35 and

chromosome 13q loci, as well as evidence for an interaction between these loci. We identified the responsible mutation in the gene on 1p35 encoding for the LDL receptor adaptor protein and described the intron-exon structure of the gene. Finally, we show evidence of linkage in normal dizygotic twin children and their parents between LDL levels and the chromosome 1p35 locus. Our findings indicate that this gene may be of relevance to the general population and also underscore the importance of the chromosome 13q locus.

### Materials and Methods

We studied a Druze family that resides in southern Syria. The Druze are a distinct ethnic group defined by descent and endogamy.<sup>9</sup> They are concentrated in well-defined areas of the middle East, primarily in Lebanon and Syria. The family consists of a core pedigree that came to our attention because 3 children developed typical features of hypercholesterolemia, including giant tendon xanthomas. Seventy-two members of the family were recruited to participate in the study after informed consent was obtained from each individual or from a legal guardian. Total cholesterol, high-density lipoprotein cholesterol (HDL), and triglycerides were measured by an automated method after a 12-hour fast. The LDL values were calculated by

Original received October 29, 2001; revision received April 1, 2002; accepted April 1, 2002.

From the Franz Volhard Clinic and Max Delbrück Center for Molecular Medicine (H.A.-K., S.B., K.H., A.B., G.N., F.C.L.), Medical Faculty of the Charité, Humboldt University of Berlin, Germany; the Department of Biochemistry (H.A.-K., M.J.), College of Pharmacy, University of Damascus, Syria; Institute of Molecular Genetics and Faculty of Medicine (H.A.-K., E.K.F.B., H.A.D.), University of Heidelberg; Institute for Medical Biometry (K.S.), Informatics and Epidemiology, University of Bonn; and Infogen GmbH (A.B.), subsidiary of Valigen NV, Berlin, Germany.

Correspondence to Friedrich C. Luft, Wiltberg Strasse 50, 13125 Berlin, Germany. E-mail luft@fvk-berlin.de

© 2002 American Heart Association, Inc.

*Circulation Research* is available at <http://www.circresaha.org>

DOI: 10.1161/01.RES.0000018002.43041.08

**TABLE 1. Characteristics of Core Pedigree Members**

Subjects	Age, y	Sex	BMI, kg/m <sup>2</sup>	Apo E	TC	TGL	LDL	HDL
4.5	58	F	27	3,3	6.4	1.6	4.5	1.16
3.3	73	M	24	3,3	4.9	1.2	3.1	1.24
5.9	28	M	19	3,3	9.6	0.9	8.6	0.59
5.10	30	F	20	3,3	14.5	1.2	13.3	0.62
5.11	26	F	19	3,3	13.3	0.9	12.25	0.69
5.12	21	F	23	3,3	4.3	1.1	2.4	1.39
5.13	23	F	24	3,3	4.0	1.0	2.7	1.08
5.14	33	F	22	3,3	5.0	0.7	3.3	1.52

Core pedigree members (numbered from the large pedigree) with demographic information, Apo E genotype status, and total cholesterol (TC), triglyceride (TGL), low-density lipoprotein cholesterol (LDL), and high-density lipoprotein cholesterol (HDL) concentrations (mmol/L).

means of the Friedewald formula.<sup>10</sup> Genotyping was done with the ABI PRISM genotyping system (Applied Biosystems) as described earlier.<sup>8</sup> All markers were amplified under a common set of PCR conditions by use of the True Allele PCR mix (PE Biosystems) containing AmpliTaq Gold DNA Polymerase and a final MgCl<sub>2</sub> concentration of 2.5 mmol/L, with a 56, 58, and 60°C annealing temperatures. Electrophoresis and detection were done on an ABI 377 equipped with Genescan 3.1 software (Applied Biosystems). Genotyping was performed with Genotyper 2.0 software (Applied Biosystems). Genotypes were exported as a text file for subsequent linkage analysis. To test the hypothesis that an interaction takes place between chromosome 1 and 13 loci, GENEHUNTER-TWOLOCUS software was used according to Strauch et al.<sup>11</sup> The program allowed us to perform parametric and nonparametric linkage analysis with 2 trait loci. In the parametric context, specific 2-locus models of disease may be tested.

An expanded Materials and Methods section can be found in the online data supplement available at <http://www.circresaha.org>.

## Results

The core pedigree consisted of 2 second-cousin parents with normal LDL values who had 6 children, as shown in Table 1. Three children had LDL values consistent with FH and had tendon xanthomas. Three other children had normal LDL values. The members of the core family stem from an extended normocholesterolemic pedigree (online Figure 1 in the online data supplement available at <http://www.circresaha.org>). We first excluded the LDL receptor locus by direct sequencing and the apolipoprotein (apo) B locus by genotyping locus-linked microsatellite markers. We also genotyped apo E in the core family. All the members were homozygous for apo E3. We performed a total genome scan of the core pedigree (online Figure 2). The affected probands in the core pedigree were homozygous for the marker D1S199, whereas the nonaffected siblings were heterozygous. Two-point linkage analysis yielded a logarithm of the odds (LOD) score of 2.37 for this marker at recombination fraction ( $\theta=0.00$ ). We genotyped the markers D1S2667, D1S186, and 12 markers in between. These markers cover a wide region that contains the tumor necrosis factor receptor superfamily member 1B (TNFRSF1B). This gene has been implicated in familial-combined hyperlipidemia.<sup>12,13</sup> To exclude TNFRSF1B, we sequenced the gene and found no mutations.

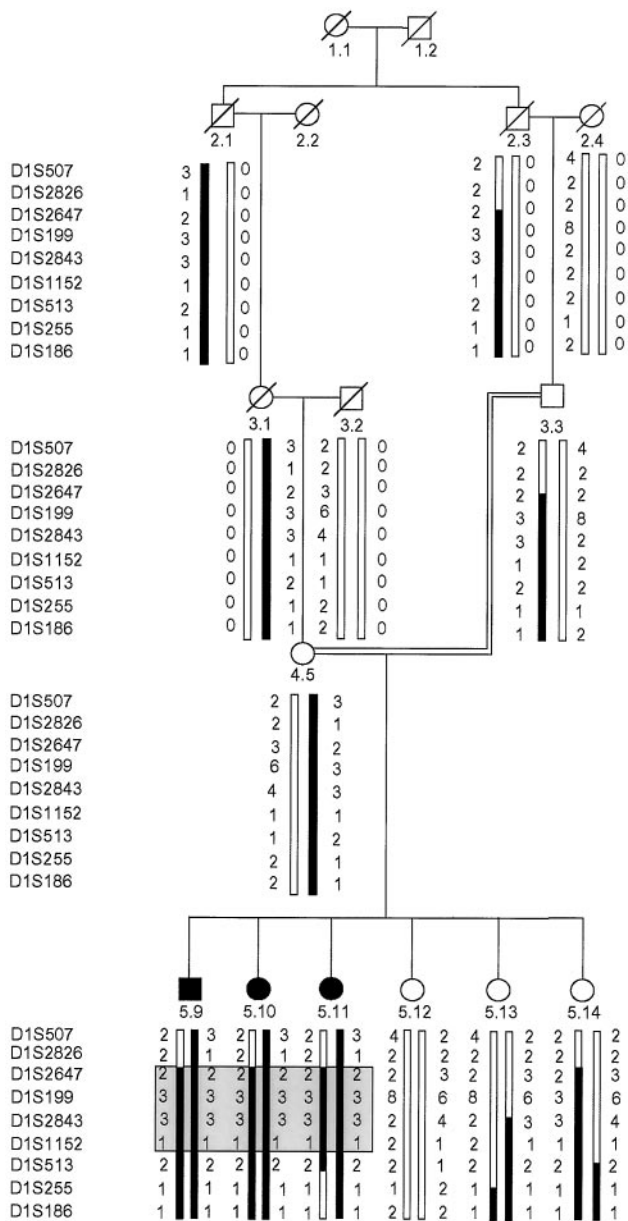
The recombination event in the individual 5.14 at D1S513 established the proximal boundary of the region as shown by the haplotypes in Figure 1. An ancestral recombination event

defined the distal boundary of the region at D1S2826. We then genotyped all the family members for the markers in the region. These marker data enabled us to estimate a more accurate allele frequency, independent of that available in the Genome Database. Table 2 shows the calculated 2-point LOD score in the core family depending on the observed allele frequency. The highest 2-point LOD score was 2.61 at the marker D1S2843 ( $\theta=0.00$ ).

Figure 2 (top panel) shows the calculated multipoint LOD scores for all genotyped markers in the region. We examined whether or not this locus contributes quantitatively (QTL) to cholesterol, LDL, HDL, and triglyceride levels. Heterozygous persons had total cholesterol, LDL, HDL, and triglyceride concentrations that were not significantly different from the 44 individuals with the wild-type haplotypes, consistent with an autosomal recessive mode of inheritance.

The genome scan yielded an additional LOD score of 2.63 at the markers D13S1267 and D13S1240. Genotyping more markers enabled us to identify a region of homozygosity on chromosome 13. The boundaries of this region are defined by the recombination events that had occurred in the individual 4.5 at the markers D13S1300 and D13S1266. Haplotypes in the core family are shown in Figure 3. The region thus extends over 13q22-q32, a 14-cM interval. We found this result remarkable in light of our earlier finding that a cholesterol-influencing gene resides at this locus on chromosome 13.<sup>8</sup> We therefore genotyped chromosome 13 markers in the entire pedigree. We found one additional homozygous person (5.15) who has the same haplotype as the affected offspring. This individual is also a daughter of the proband's father, but stems from a different mother. Table 2 gives the 2-point LOD scores. The multipoint analysis yielded a LOD score of 3.08 shown in Figure 2 (bottom panel). We tested this locus as a QTL in the family; however, the locus did not have a QTL effect.

We had observed in our earlier study that the chromosome 13 locus was linked to total and LDL cholesterol in a panel of dizygotic twins and their parents from Berlin.<sup>8</sup> We therefore also genotyped this group for the markers D1S552, D1S199, and D1S2843. This analysis yielded linkage of total and LDL cholesterol to the chromosome 1 locus. The multipoint LOD score for total cholesterol was 1.82 ( $P=0.0019$ ) and 1.86 ( $P=0.0017$ ) for LDL cholesterol.



**Figure 1.** Haplotype segregation in the core pedigree from the candidate region on chromosome 1. Black, Haplotypes associated with the disease; White, haplotypes not associated with the disease. Ancestral haplotypes are reconstructed according to those of the probands. Solid symbols are affected homozygous persons. Double lines indicate consanguineous marriages. The gray bar shows the region of homozygosity that cosegregates with the disease in the affected individuals. Numbers refer to those in the extended pedigree. Region of homozygosity extends over the interval between the flanking markers D1S2826 D1S513, an 18-cM interval.

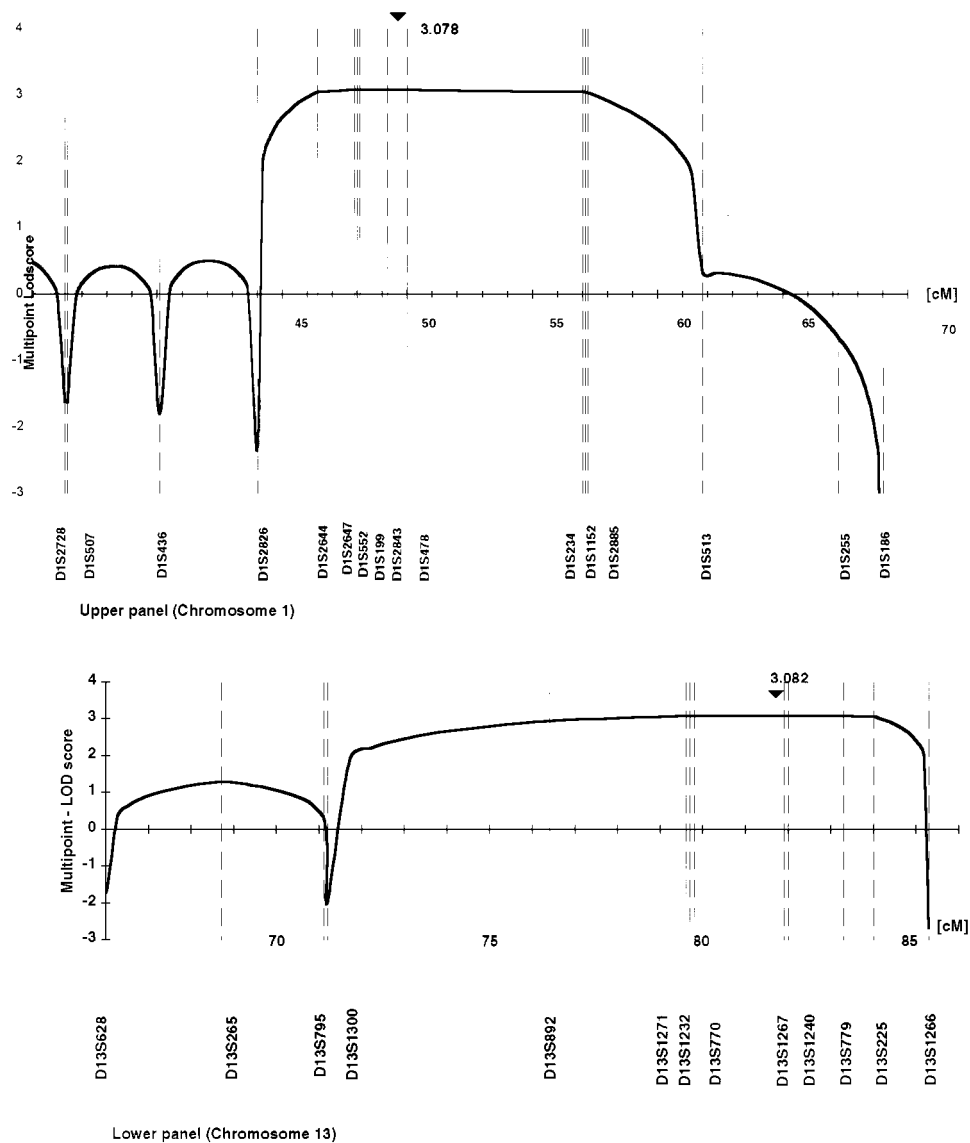
In the single-locus analysis, the FH trait appeared to be recessive at both the chromosome 1 and 13 loci. All 3 affected children in the core family (individuals 5.9, 5.10, and 5.11) were homozygous by descent for a few markers at both regions. We regarded this remarkable result as a clear incentive to perform a linkage analysis with 2-trait loci. We used GENHUNTER-TWOLOCUS to test the hypothesis that only a simultaneous homozygous-mutant genotype at the trait loci on both chromosome 1 and 13 leads to the disease

**TABLE 2.** Two-Point LOD Scores in the Core Pedigree on Chromosome 1 and 13

Marker	Recombination Fraction $\theta$						
	0.0	0.01	0.05	0.1	0.2	0.3	0.4
D1S507	-1.629	-0.333	0.229	0.374	0.357	0.211	0.050
D1S2826	-1.58	-0.36	0.20	0.35	0.33	0.19	0.04
D1S199	2.37	2.32	2.11	1.85	1.30	0.75	0.26
D1S2843	2.61	2.55	2.31	2.01	1.41	0.83	0.31
D1S513	0.675	0.658	0.589	0.503	0.336	0.184	0.064
D1S186	$\infty$	-1.560	-0.327	0.068	0.256	0.199	0.070
D13S628	-1.45	-0.75	-0.17	0.03	0.11	0.03	-0.04
D13S1300	-0.88	0.01	0.64	0.84	0.82	0.60	0.28
D13S795	1.94	1.90	1.73	1.52	1.07	0.61	0.19
D13S892	2.38	2.34	2.18	1.97	1.50	1.00	0.47
D13S1232	2.22	2.18	2.02	1.82	1.39	0.93	0.45
D13S770	2.42	2.36	2.15	1.87	1.33	0.81	0.34
D13S1267	2.63	2.58	2.34	2.05	1.45	0.88	0.38
D13S1240	2.63	2.57	2.35	2.07	1.50	0.92	0.38
D13S225	2.30	2.25	2.03	1.75	1.19	0.65	0.21
D13S1266	-1.88	-0.20	0.38	0.53	0.52	0.38	0.19

(“recessive and recessive”), a multiplicative 2-locus model.<sup>14</sup> In addition to the nonparametric analysis with 2 disease loci, we performed a parametric 2-locus analysis under the assumption of this multiplicative model. We also analyzed the data under a heterogeneity model and under an additive model assuming a recessive action of both loci. The models are shown in Table 3. Because 2-locus analysis is computationally much more intensive than an analysis with only 1 disease locus, we were only able to include 4 of the 6 children of the core family. We first chose to include the 3 affected children and 1 unaffected child (5.13) in the GENHUNTER-TWOLOCUS analysis. The nonparametric 2-locus analysis yielded a maximum nonparametric linkage (NPL) score of 8.64 ( $P=3.8 \times 10^{-6}$ ). The  $P$  value reflects a test of the null hypothesis that both disease loci are unlinked to their corresponding marker maps. The parametric analysis yielded a maximum 2-locus LOD score of 5.41 under the multiplicative model. The heterogeneity and additive models yielded maximum LOD scores of 2.82 and 2.75, respectively. For nonparametric analysis, as well as parametric analysis under all 3 models, the maximum score occurred directly at the markers D1S2843 and D13S1267, the same position for the maximal single-locus LOD score. The results remained the same when we replaced individual 5.13 by another unaffected child; the maximal change in the LOD-score was less than 1%. To compare the results of the parametric 2-trait locus analysis to the single-locus results, we next recalculated the single-locus LOD scores for the pedigree without the individuals 5.12 and 5.14. The maximum LOD scores were 2.80 for chromosome 1 and 2.83 for chromosome 13. Hence, the 2-locus LOD score for the multiplicative model was superior to both single-locus LOD scores.

The superiority of the multiplicative model, compared with the other models, could be because of the fact that 3 of the 4



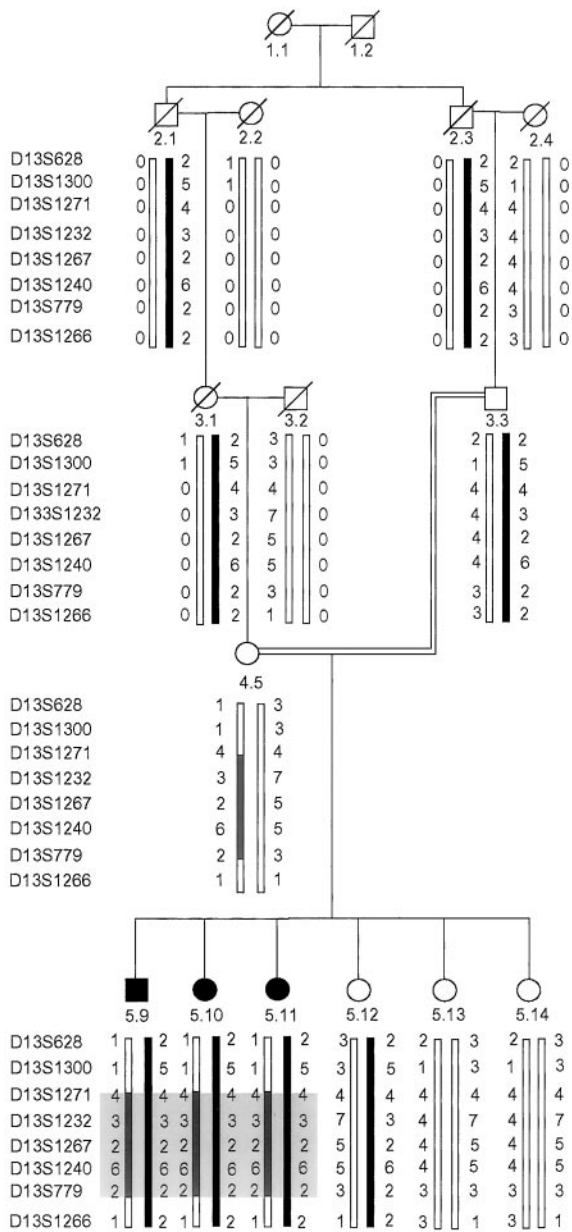
**Figure 2.** Multipoint linkage analyses of ARH in the entire family. All available individuals from the entire family (see online Figure 1) were genotyped for the markers in both linked regions on chromosome 1 and 13. Top, Multipoint LOD scores in the entire family on chromosome 1 locus; the highest LOD score was obtained at the marker D1S2843, 3.078. Bottom, Multipoint LOD scores in the entire family on chromosome 13 locus; the highest LOD score was obtained at the marker D13S1267, 3.082. The program SimWalk was used for computation of multipoint LOD scores. Linkage analyses were performed assuming an autosomal-recessive mode of inheritance with a complete penetrance and a disease allele frequency of 0.0001. Markers are indicated below the x-axis.

children were affected. To test this possibility, we repeated the 2-trait locus analysis using all 3 unaffected children, but only 1 affected child 5.9. The resulting LOD scores were 3.06 for the multiplicative model, 1.86 for the heterogeneity model, and 1.66 for the additive model. Again, the LOD scores did not change by more than 1% when another affected child was selected for the analysis.

We identified the genomic structure of ARH as shown in Table 4. First, we scanned the bacterial artificial chromosome (BAC) library with primers from the cDNA of ARH and found that the 121-O3 clone harbored the ARH gene. ARH-specific primers, designed based on the ARH cDNA sequence, were used for sequencing the cloned genomic DNA (BAC clone) to determine the positions of the introns and the intron-exon boundaries given in Table 4. Intron-exon bound-

aries were defined by comparing the genomic sequence with the cDNA sequence. ARH consists of 9 exons. The exon size ranges from 35 (exon 8) to 2066 bp (exon 9, the last exon). Exon 1 contains 21 bp of untranslated sequence followed by the initiation codon ATG. All the exon-intron boundary sequences conform to the consensus splice donor (GT) and acceptor (AG) sites, with the exception of the splice donor site of intron 8, which is GG instead of GT.

We sequenced the ARH gene in the core family and found a mutation that lies in the acceptor site of intron 1. The mutation converts the original acceptor-splice site from G→C, a transversion mutation. Figure 4A shows the intron 1/exon 2 junction in a normal proband (top) and in an affected child (bottom). The affected offspring are all homozygous for the mutation, whereas the parents are heterozygous. Figure



**Figure 3.** Haplotype segregation in the core pedigree from the candidate region on chromosome 13 is given. Black, Haplotypes associated with the disease; White, haplotypes not associated with the disease. Solid symbols are affected homozygous persons. Double lines indicate consanguineous marriages. The gray bar shows the region of homozygosity in the affected individuals. Numbers refer to those in the extended pedigree. The region of homozygosity extends over the interval between the flanking markers D13S1300 and D13S1266, a 14-cM interval.

4B shows the mutated nucleotide sequence of the coding and noncoding strands at the intron 1/exon 2 junction. We found a restriction enzyme *PvuII* to cut the PCR product of exon 2 at the mutated donor-acceptor site. We used this enzyme to screen the entire family by using the exon 2 PCR product. On recognition, the enzyme cuts the 316-bp PCR product of exon 2 producing 2 fragments, 256 and 60 bp, in the homozygous state and 3 fragments, 316, 256, and 60 bp, in the heterozygous state. Figure 4C shows the status of the core family members with regard to the ARH G transversion mutation.

**TABLE 3.** Two-Locus Disease Models Studied for the Chromosome 1 and 13 Data of Familial Hypercholesterolemia

Chr. 1 Locus Genotype	Chr. 13 Locus Genotype		
	+/+	Het	m/m
<b>Multiplicative model</b>			
+/+	0	0	0
Het	0	0	0
m/m	0	0	1
<b>Heterogeneity model</b>			
+/+	0	0	1
Het	0	0	1
m/m	1	1	1
<b>Additive model</b>			
+/+	0	0	0.5
Het	0	0	0.5
m/m	0.5	0.5	1

Penetrance is given for 2-locus genotype combinations, with the genotype for the locus on chromosome 1 shown at the left and on the top for chromosome 13. Multiplicative, heterogeneity, and additive model are shown, with the assumption of a recessive action of both genes for all 3 models. The disease allele frequency was set to 0.0001 for both loci. +/+ indicates wild-type genotype; Het, heterozygous genotype; and m/m, mutant genotype.

The parents (4.5 and 3.3) are heterozygous for the mutation. The 3 affected children (5.9, 5.10, and 5.11) are homozygous. In agreement with the reported haplotypes, 2 of unaffected children (5.13 and 5.14) are carriers and 1 (5.12) is a noncarrier. We screened the entire family and found 21 individuals who were heterozygous for the mutation. These individuals are the same who share the haplotype with the affected patients for the microsatellite markers genotyped in the region. Only the 3 affected offspring in the core family are homozygous for the mutation. We also screened 31 normal Syrian individuals for the mutation and another 30 normal unrelated German individuals. None had the mutation.

We also found 3 single nucleotide polymorphisms (SNPs) in ARH gene. One variant lies in exon 6 at bp 625 of ARH cDNA converting C→T and leads to an amino acid substitution from Pro to Ser. The second SNP lies in the intron 6 at bp 28 converting A→G. The third SNP lies in exon 7 at bp 675 of ARH cDNA converting A→G but without any change in the coding sequence. To explore a possible effect of the variant in exon 6 on the phenotype, we genotyped the whole family for these 3 SNPs. Five persons were homozygous for the variant in exon 6. We also genotyped our twin panel for the 3 SNPs. The data showed no significant association between these variants and LDL levels ( $P < 0.1$ ).

**Discussion**

Our linkage data provide strong evidence that ARH is influenced by 2 loci in the family in our present study, namely 1p36.1-p35 and 13q22-q32. Our findings support the recent observations reported by Eden et al.<sup>6</sup> Furthermore, they corroborate earlier findings from our group regarding chromosome 13 as a locus for a gene influencing cholesterol concentrations.<sup>8</sup> We also excluded the recently described autosomal-dominant FH locus on chromosome 1 and the

**TABLE 4. Genomic Structure of the ARH Gene**

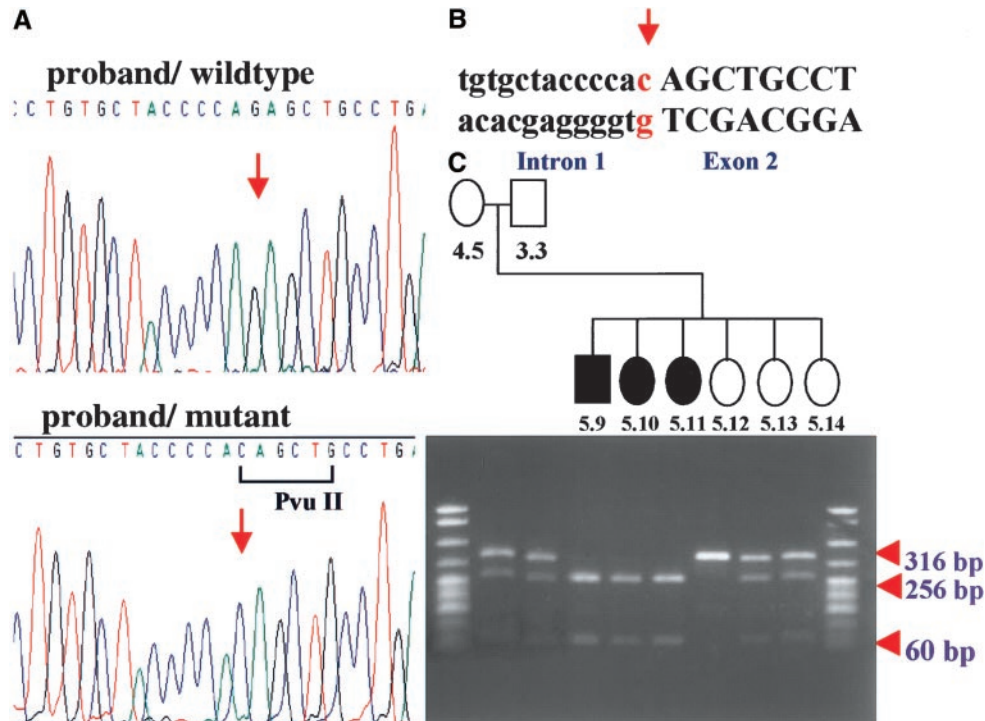
Exon No.	Exon Length, bp	cDNA Position	Splice Donor	Splice Acceptor	Intron No.	Interrupted Amino Acid
1	109	1–109	ACCGCAgtgagt ...	...cccagAGCTGC	1	L*ys
2	143	110–252	GCTACAgtgagc ...	...cccaagGCTAAG	2	-
3	113	253–365	ATACAGgttacgc ...	...tttcagGATCTC	3	Ar*g
4	115	366–480	AAGATGgtcagc ...	...tttcagGCACAG	4	-
5	73	481–553	AGGAAGgtgaga ...	...ctgcagAGAAAG	5	G*lu
6	84	554–637	AGAGCTgtgagt ...	...tcacagTGGTCG	6	L*eu
7	131	638–768	GTCTGGgtgagt ...	...cccaagGAGCTG	7	-
8	35	769–803	TTCGAGggaatg ...	...ttccagGCTTGC	8	Ar*g
9	2066	804–2869				

ARH-1 locus on chromosome 15 mapped by Varret et al<sup>3</sup> and Ciccarese et al,<sup>5</sup> respectively.

In addition to the chromosome 1p36.1-p35 locus, we also found compelling evidence for linkage to 13q22-q32. We do not believe that this finding is related to chance. In our earlier study, we identified this locus as the site for a gene that lowers LDL.<sup>8</sup> We studied a family with autosomal-dominant familial hypercholesterolemia that had affected heterozygous and homozygous members with lower than expected LDL values. We corroborated this locus with a linkage analysis in dizygotic twins and their parents. Eden et al<sup>6</sup> did not report any evidence of linkage to 13q22-q32 in their study. How-

ever, the family in our present study and the family we reported earlier are both unrelated Arab families<sup>8</sup>; their genetic background may be similar.

Garcia et al<sup>7</sup> cloned the gene responsible for ARH residing at chromosome 1p35. Their exciting finding was a new protein that is responsible for LDL receptor malfunction. The gene they identified codes for a protein that has a phosphotyrosine binding domain PTD. This domain was found to interact with the cytoplasmic tail of the LDL receptor.<sup>15</sup> Malfunction of this interaction could account for a disabled LDL receptor. ARH is expressed in many tissues, even in those that have only a very low LDL



**Figure 4.** Detection of the intron 1 acceptor-site mutation in the core family. A, Sequence analysis of the intron 1/exon 2 junction in a normal proband (top) and in an affected child (bottom). The nucleotide mutated in the affected child is indicated (arrowhead). B, Nucleotide sequence of the coding and noncoding strands at the intron 1/exon 2 junction. The mutation affects the variant guanine of the acceptor-splice site converting it into cytosine. C, Status of family members with regard to the ARHG transversion mutation. The 316-bp exon 2 PCR fragment encompassing the intron 1/exon2 junction was digested with *PvuII* enzyme, which recognizes the mutated sequence at its restriction site CAGCTG. The parents (4.5 and 3.3) are heterozygous for the mutation and the 3 affected children (5.9, 5.10, and 5.11) are homozygous. In agreement with the reported haplotypes, two (5.13 and 5.14) of the unaffected children are carriers and one (5.12) is noncarrier.

receptor expression. Conceivably ARH has other functions; however, similar to the families described by Garcia et al,<sup>7</sup> our affected family members have no other discernible phenotypes.

As reported in the present study, we determined the exon-intron structure of ARH. All splice sites conform to the GT/AG rule, except for intron 8, where the donor splice site GT was converted to GG. However, there are a few exceptions to this rule.<sup>16–20</sup> We identified the intron 1 acceptor-site mutation in the affected persons of the core family. The mutation replaces the G bp of the acceptor-splice site (AG) by C. We corroborated our findings by analysis of the other family members. We also examined 31 healthy Syrians and 30 Germans. All were wild-type for the mutation. A defect in intron 1 splice-donor site will either skip exon 2 on splicing or will retain intron 1 in the processed mRNA. Both cases contain stop codons that would result in a truncated protein missing the amino acids coded by exons 2 to 9. We also found 3 SNPs in exon 6, intron 6, and exon 7. The SNP in exon 6 causes an amino acid exchange Pro→Ser. We excluded this variant when we found other homozygous family members who had normal total cholesterol and LDL levels. Moreover, we examined these variants in our twin panel and found that they were not associated with total cholesterol and LDL levels.

The parametric and nonparametric 2-locus analysis was not only done to verify that ARH is indeed governed by 2 loci, but also to gain insight into how the 2 genes may act together. Due to the restrictions on the pedigree size with GENEHUNTER-TWOLOCUS, we were not able to include all 6 children in the analysis simultaneously. We performed the analysis with all 3 affected children and 1 unaffected child. A recessive-and-recessive multiplicative model gave the strongest results, irrespective of the unaffected child selected. The model reflects that only individuals with a homozygous mutant genotype at both loci will develop the trait. Here, the 2-locus LOD score was 5.41, which was almost twice as high as the result obtained for the heterogeneity and additive models. With the latter models, the disease may be caused by a homozygous-mutant genotype at only 1 locus.

It may be argued that the multiplicative 2-locus model gave the strongest results merely because 2 unaffected individuals were not included in the analysis. We therefore repeated the 2-locus calculations with all 3 unaffected children and 1 affected child. The resulting LOD score was 3.06 for the multiplicative model, 1.86 for the heterogeneity model, and 1.66 for the additive model. Hence, even in this setting the superiority of the multiplicative model remained over the other 2 models. These results support a multiplicative action of both loci, where a homozygous mutant genotype at both chromosome 1 and 13 loci is necessary to express the disease.

Our 2-locus multiplicative model suggests that a mutation at only 1 locus is not sufficient to develop the disease. However, our results suggest that the chromosome 1 mutation has a dramatic effect on the adaptor protein. Our analysis raises the possibility that 2 proteins coded by 2 separate genes at distant loci are able to perform the same biological function. A mutation at 1 locus may disable one of the

proteins, whereas the second is still available for normal function. Such a model would require disruption of both genes, disabling both proteins to express the phenotype. If our hypothesis is correct, a mutation on chromosome 1 leading to a dramatically altered protein is still in accord with the multiplicative 2-locus trait model. Our linkage and interaction analysis strongly supports the hypothesis that 2 loci are involved in this form of recessive FH in our family.

We were interested to observe linkage for the 1p36.1-p35 locus to total and LDL cholesterol in normal dizygotic twins and their parents. We verified earlier findings that LDL, HDL, total cholesterol, triglycerides, and body mass index are all strongly influenced by genetic variance.<sup>21</sup> In the Syrian family, the chromosome 1 locus did not appear as a QTL. The effect here is autosomal-recessive and qualitative in nature. In the twins, the locus appeared as a QTL, suggesting that other variations in the responsible gene are operative. We believe that these studies are unique because we not only mapped a digenic syndrome, but also suggested its relevance in a completely unrelated, normal population. Identification of ARH and its mutations are clearly important to our understanding of cholesterol metabolism and cardiovascular risk. Elucidating the gene at the chromosome 13q locus and the basis for the interaction will have a high priority.

### Acknowledgments

Hussam Al-Kateb is supported by a fellowship from the Deutscher Akademischer Austauschdienst (DAAD). We thank Atakan Aydin for his help in solving technical problems, and Athanasios Vergopolous for sequencing the LDL receptor in the family.

### References

- Goldstein JL, Brown MS. Familial hypercholesterolemia: identification of a defect in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity associated with overproduction of cholesterol. *Proc Natl Acad Sci U S A*. 1973;70:2804–2808.
- Myant NB. Familial defective apolipoprotein B-100: a review, including some comparisons with familial hypercholesterolaemia. *Atherosclerosis*. 1993;104:1–18.
- Varret M, Rabes JP, Saint-Jore B, Cenarro A, Marinoni JC, Civeira F, Devillers M, Krempf M, Coulon M, Thiar R, Kotze MJ, Schmidt H, Buzzi JC, Kostner GM, Bertolini S, Pocovi M, Rosa A, Farnier M, Martinez M, Junien C, Boileau C. A third major locus for autosomal dominant hypercholesterolemia maps to 1p34.1-p32. *Am J Hum Genet*. 1999;64:1378–1387.
- Hunt SC, Hopkins PN, Bulka K, McDermott MT, Thorne TL, Wardell BB, Bowen BR, Ballinger DG, Skolnick MH, Samuels ME. Genetic localization to chromosome 1p32 of the third locus for familial hypercholesterolemia in a Utah kindred. *Arterioscler Thromb Vasc Biol*. 2000;20:1089–1093.
- Ciccarese M, Pacifico A, Tonolo G, Pintus P, Nikoshkov A, Zuliani G, Fellin R, Luthman H, Maioli M. A new locus for autosomal recessive hypercholesterolemia maps to human chromosome 15q25-q26. *Am J Hum Genet*. 2000;66:453–460.
- Eden ER, Naoumova RP, Burden JJ, McCarthy MI, Soutar AK. Use of homozygosity mapping to identify a region on chromosome 1 bearing a defective gene that causes autosomal recessive homozygous hypercholesterolemia in two unrelated families. *Am J Hum Genet*. 2001;68:653–660.
- Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, Calandra S, Bertolini S, Cossu F, Grishin N, Barnes R, Cohen JC, Hobbs HH. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science*. 2001;292:1394–1398.
- Knoblauch H, Müller-Myhsok B, Busjahn A, Ben Avi L, Bähring S, Baron H, Heath SC, Regina Uhlmann, Faulhaber H-D, Shpitzen S, Aydin A, Reshef A, Rosenthal M, Eliav O, Mühl A, Lowe A, Schurr D, Harats D, Jeschke E, Friedlander Y, Schuster H, Luft FC, Leitersdorf E. A

- cholesterol-lowering gene exists on chromosome 13q. *Am J Hum Genet.* 2000;66:157–166.
9. Nevo S. Genetic blood markers in the Arab Druze of Israel. *Am J Phys Anthropol.* 1988;77:183–190.
  10. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem.* 1972;18:499–502.
  11. Strauch K, Fimmers R, Kurz T, Deichmann KA, Wienker TF, Baur MP. Parametric and nonparametric multipoint linkage analysis with imprinting and two-locus-trait models: application to mite sensitization. *Am J Hum Genet.* 2000;66:1945–1957.
  12. Glenn CL, Wang WY, Benjafield AV, Morris BJ. Linkage and association of tumor necrosis factor receptor 2 locus with hypertension, hypercholesterolemia and plasma shed receptor. *Hum Mol Genet.* 2000;9:1943–1949.
  13. Geurts JM, Janssen RG, van Greevenbroek MM, van Der Kallen CJ, Cantor RM, Bu Xd, Aouizerat BE, Allayee H, Rotter JI, de Bruin TW. Identification of TNFRSF1B as a novel modifier gene in familial combined hyperlipidemia. *Hum Mol Genet.* 2000;9:2067–2074.
  14. Risch N. Linkage strategies for genetically complex traits, I: multilocus models. *Am J Hum Genet.* 1990;46:222–228.
  15. Trommsdorff M, Borg JP, Margolis B, Herz J. Interaction of cytosolic adaptor proteins with neuronal apolipoprotein E receptors and the amyloid precursor protein. *J Biol Chem.* 1998;273:33556–33560.
  16. Green MR. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu Rev Cell Biol* 1991;7:559–599.
  17. Dong J, Kukula AK, Toyoshima M, Nakajima M. Genomic organization and chromosome localization of the newly identified human heparanase gene. *Gene.* 2000;253:171–178.
  18. Broderick TP, Schaff DA, Bertino AM, Dush MK, Tischfield JA, Stambrook PJ. Comparative anatomy of the human APRT gene and enzyme: nucleotide sequence divergence and conservation of a non-random CpG dinucleotide arrangement. *Proc Natl Acad Sci U S A.* 1987;84:3349–3353.
  19. Schroeder HW Jr, Walter MA, Hofker MH, Ebens A, Willems van Dijk K, Liao LC, Cox DW, Milner EC, Perlmutter RM. Physical linkage of a human immunoglobulin heavy chain variable region gene segment to diversity and joining region elements. *Proc Natl Acad Sci U S A.* 1988;85:8196–8200.
  20. Alexandre D, Chuchana P, Brockly F, Blancher A, Lefranc G, Lefranc MP. First genomic sequence of a human Ig variable lambda gene belonging to subgroup I: functional genes, pseudogenes and vestigial sequences are interspersed in the IGLV locus. *Nucleic Acids Res.* 1989;17:3975.
  21. Knoblauch H, Busjahn A, Münter S, Nagy Z, Faulhaber H-D, Schuster H, Luft FC. Heritability analysis of lipids and three gene loci in twins link the macrophage scavenger receptor to high-density lipoprotein cholesterol concentrations. *Arterioscler Thromb Vasc Biol.* 1997;17:2054–2060.



# Circulation Research

JOURNAL OF THE AMERICAN HEART ASSOCIATION



## Mutation in the ARH Gene and a Chromosome 13q Locus Influence Cholesterol Levels in a New Form of Digenic-Recessive Familial Hypercholesterolemia

Hussam Al-Kateb, Sylvia Bähring, Katrin Hoffmann, Konstantin Strauch, Andreas Busjahn, Gudrun Nürnberg, Muhidien Jouma, Eckehard K.F. Bautz, Hans A. Dresel and Friedrich C. Luft

*Circ Res.* 2002;90:951-958; originally published online April 11, 2002;  
doi: 10.1161/01.RES.0000018002.43041.08

*Circulation Research* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231  
Copyright © 2002 American Heart Association, Inc. All rights reserved.  
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://circres.ahajournals.org/content/90/9/951>

Data Supplement (unedited) at:

<http://circres.ahajournals.org/content/suppl/2002/05/10/90.9.951.DC1>

**Permissions:** Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation Research* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

**Reprints:** Information about reprints can be found online at:  
<http://www.lww.com/reprints>

**Subscriptions:** Information about subscribing to *Circulation Research* is online at:  
<http://circres.ahajournals.org/subscriptions/>

## EXPANDED METHODS

Two point LOD-score calculations were performed by the LINKAGE program package using the LINKRUN computer program (Prof. T.F. Wienker, Bonn, Germany, personal communication). We assumed an autosomal recessive mode of inheritance with a complete penetrance and a disease allele frequency of 0.0001. Marker allele frequency was calculated from the observed allele frequencies in unrelated persons from the same Druze population (recruited mainly from the spouses in-law within the extended pedigree). The program SIMWALK was used for computation of multipoint LOD scores.<sup>1</sup> Haplotyping was performed with GENHUNTER version 2.0 and by hand.<sup>2-3</sup> For two-locus analysis, we used the program GENEHUNTER-TWOLOCUS that allowed us to perform parametric and nonparametric analysis.<sup>4</sup> The two trait loci should be unlinked (e.g. located on different chromosomes), and each of them is assumed to be linked to a separate marker map. For two-locus LOD and NPL-score calculation, GENEHUNTER-TWOLOCUS simultaneously uses all available markers. Hence, the approach is completely multipoint, and a maximum of linkage information is extracted at both genetic regions under study. QTL effects for the traits, total cholesterol, LDL, HDL, and triglycerides were tested by applying the TDT statistic for quantitative traits as described by Allison and as implemented in QTDT.<sup>5</sup> Significance was ascertained by means of 1000 Monte-Carlo permutations.

Using GENEHUNTER-TWOLOCUS, we relied on three models, a multiplicative model, a heterogeneity model, and an additive model. With a heterogeneity model, a disease-predisposing genotype at either locus per se can cause the trait. With an additive model, the penetrance for predisposing genotypes at both loci equals the sum of the penetrances for individuals having an at-risk genotype at one locus but not the other. The multiplicative model reflects epistasis or interaction. Strictly speaking, this also holds for the heterogeneity model. Here, the two-locus effect differs from the sum of the two single-locus effects, just as it does for the multiplicative model. We

used the same marker maps as for the single-locus multipoint analyses for all models. A disease allele frequency of 0.0001 was assumed at both trait loci.

To test the hypothesis that the ARH locus on chromosome 1 is of relevance to the general population, we performed studies in 100 pairs of dizygotic (DZ) twins and available parents of the DZ twins. The subjects were all healthy, normotensive Caucasians recruited from various parts of Germany. The protocol was approved by the Humboldt University's committee on the protection of human subjects and written informed consent was obtained from all participants. Persons with histories of familial lipid disorders were excluded. Details on our twin analysis have been published elsewhere.<sup>6</sup>

### **Determining of gene structure of ARH**

A human bacterial artificial chromosome (BAC) DNA pools (Research Genetics) were screened by PCR with primers 5' CATGGACGCGCTCAAGTC 3' and 5' GCGTGTCTGTCCAGTTCTCA that were designed based on the cDNA sequence of the ARH gene. The first primer identified the clone 121-O3. The DNA from this clone was recovered using a standard plasmid isolation protocol (Qiagen). After determining the positions of introns and obtaining the intron-exons sequences, the following primers were used for sequencing the ARH gene in the family. Exon1: 5'CCGGGCGGAAAGTTTTT 3'and 5'CGAGTAGGTGCACAACTTGA 3', exon 2: 5'AAGAAGGCTGGTGAGAGCTG 3' and 5'CTCAGACACACC CAACTGGA, exon 3: 5'TGCCTGACAGATAGCTGGTG 3'and 5'GCCTGG ATTCAGATGGGATA, exon 4: 5'TGGAGTGGGAATAGCAGGTT 3'and 5'ACTTGGGGTCCCAGAGATT 3', exon 5: GGCTGTTCCCATTTTTTCAGA 3', and 5'GGAGAACCCCTCCCCATC 3', exon 6: 5'AAACTGCCCCTTGAGGTTG 3' and 5'TGAAAACCAAGAGGCTCTCC 3',exon 7: 5'ACATCAGAGGGGAGGGTCA 3'and 5'CCTCAGGTTCAAAGGGACA 3', exon 8: 5' CCCTGAGCTTGTGTCCTGA 3'and

5'CTAGCATTCCCTCGA AAACG 3' and exon 9: 5'GGGCACAGAGTGGGTGCT 3'. PCR was performed in 15- $\mu$ l volumes containing 1.5  $\mu$ l of 10 $\times$  PCR buffer (Eppendorf), 5 pM of each primer, 0.06 mM each dNTP, 0.06 U Taq DNA Polymerase (Eppendorf), 2.5 mM MgCl<sub>2</sub> and 100 ng total human genomic DNA. Reactions were amplified for one cycle at 95°C (4 min, hot start); 35 cycles at 94°C (30 s), (30 s) at the annealing temperature (60°C for exon 9, 62°C for exons 4 8, and 64°C for exons 2 and 3) and extension at 72°C (30s)and one cycle at 72°C (7 min). Exon 1 was amplified using the same reaction volumes as described for the other exons but in addition to 5% of DMSO and Taq Gold DNA Polymerase was instead used. Amplification step for exon 1 was performed for one cycle at 95°C (10 min); 45 cycles at 94°C (30 s), (30s) at the annealing temperature (52°C), and extension at 72°C (30 s), and one cycle at 72°C (7 min). Sequencing was done as follows: a holding step at 96°C (10 s), an annealing step for 10 s and then an extension step at 60°C for 90 s.

#### REFERENCES FOR EXPANDED METHODS

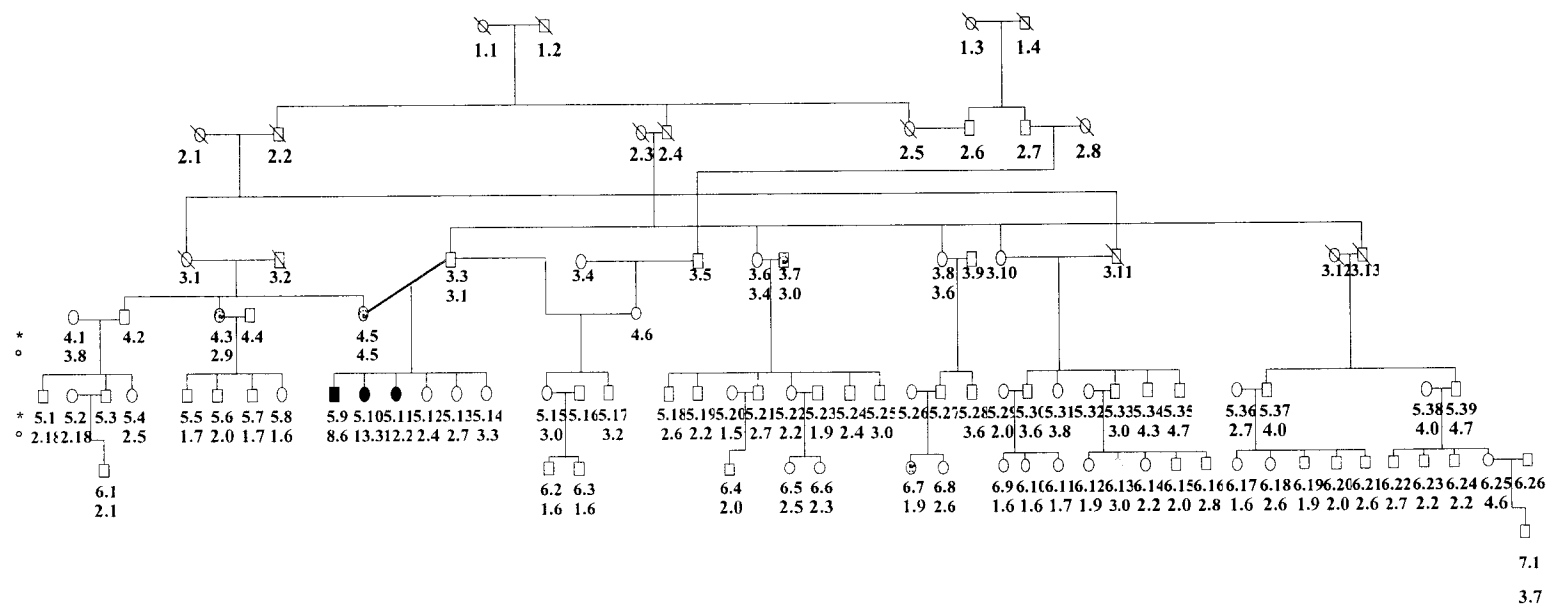
1. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am J Hum Genet.* 1996; 58: 3123-1337.
2. Kruglyak L, Daly J, Reeve-Daly P, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.* 1996; 58: 1347-1363.
3. Kruglyak L, Lander ES Faster multipoint linkage analysis using Fourier transforms. *J Comput Biol.* 1998; 5: 1-7.
4. Strauch K, Fimmers R, Kurz T, Deichmann KA, Wienker TF, Baur MP. Parametric and nonparametric multipoint linkage analysis with imprinting and two-locus-trait models: application to mite sensitization. *Am J Hum Genet.* 2000; 66: 1945-57.
5. Allison DB. Transmission-disequilibrium tests for quantitative traits. *Am J Hum Genet.* 1997; 60: 676-90.

6. Knoblauch H, Müller-Myhsok B, Busjahn A, Ben Avi L, Bähring S, Baron H, Heath SC, Regina Uhlmann, Faulhaber HD, Shpitzen S, Aydin A, Reshef A, Rosenthal M, Eliav O, Mühl A, Lowe A, Schurr D, Harats D, Jeschke E, Friedlander Y, Schuster H, Luft FC, Leitersdorf E. A cholesterol-lowering gene exists on chromosome 13q. *Am J Hum Genet.* 2000; 66: 157-166.

### **Figure legends**

(Supplement data. Fig. 1). Extended pedigree of the entire family. Solid symbols are affected persons with autosomal recessive hypercholesterolemia. Seventy-two members were examined and blood samples were obtained. The subject identification number is given and fasting low-density cholesterol (LDL-cholesterol) levels are shown.

(Supplement data. Fig.2.). Combined LOD scores resulted from a genomewide scan in the core family. A genomewide scan was performed in the core family with markers located at ~8-cM intervals. The LINKRUN computer program was used for computation of two-point LOD scores. Linkage analyses were done assuming an autosomal recessive mode of inheritance with a complete penetrance and a disease allele frequency of 0.0001. The genomewide scan in the core family resulted in two LOD scores over 2.5 on chromosomes 1 and 13.



**Figure. 1 (Supplementary data)**

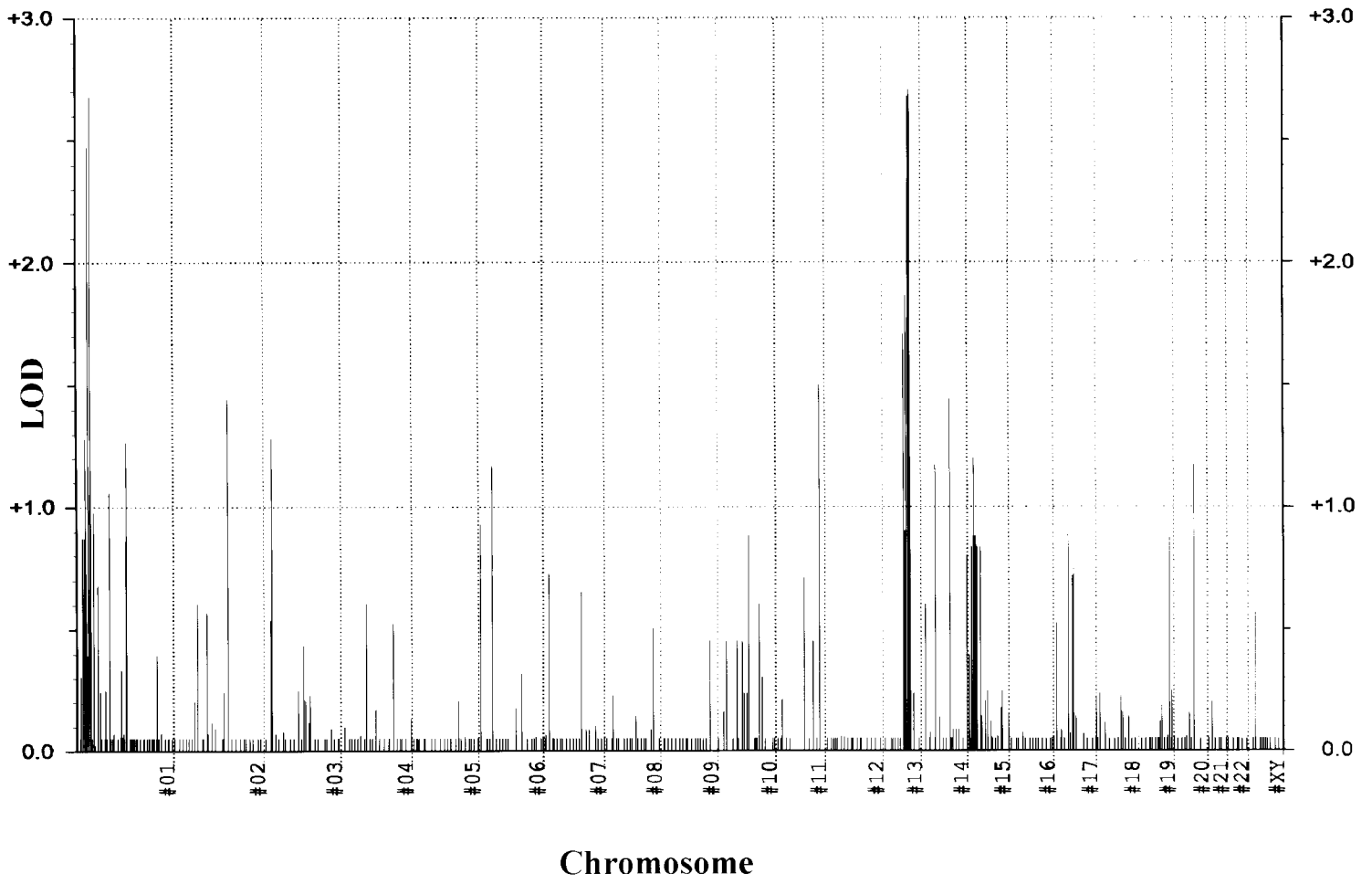


Figure.2 (Supplementary data)