Am J Hum Genet 39:300–306, 1986

A DNA Probe for the LDL Receptor Gene Is Tightly Linked to Hypercholesterolemia in a Pedigree with Early Coronary Disease

MARK F. LEPPERT,[1] SANDRA J. HASSTEDT,[2] THOMAS HOLM,[1] PETER O'CONNELL,[1] LILY WU,[3] OWEN ASH,[3] ROGER R. WILLIAMS,[4] AND RAY WHITE[1,2]

SUMMARY

A large, multigenerational family with dominantly inherited hypercholesterolemia was analyzed for genetic linkage between blood levels of low-density lipoprotein (LDL) cholesterol and the locus for the LDL receptor. A genetic marker was identified by restriction fragment length polymorphism (RFLP) in a cloned segment of the LDL receptor gene. We found no exceptions to segregation of the high-LDL cholesterol phenotype with a unique allele at the LDL receptor locus in this pedigree; tight linkage was indicated by a maximum lod score of 7.52 at $\theta = 0$. Knowledge of the LDL receptor genotype will enable investigators to study variability of phenotypic expression in response to environmental influences or to different genetic determinants.

INTRODUCTION

Genetic markers based on DNA sequence polymorphisms represent one of the most important recent developments in analysis of human genetic disease, because individual genotypes can now be defined by linkage analysis in families suspected of segregating a mutant allele of a gene of interest. When, as sometimes happens, DNA sequence polymorphisms are detected by a cloned human

Received December 12, 1985; revised April 11, 1986.

This work was supported in part by grant HL21088-09 from the National Heart, Lung and Blood Institute. R. W. is an investigator of the Howard Hughes Medical Institute.

[1] Howard Hughes Medical Institute, University of Utah Medical Center, Salt Lake City, UT 84132.

[2] Department of Human Genetics, University of Utah Medical Center.

[3] Department of Pathology, University of Utah Medical Center.

[4] Cardiovascular Genetics Research Clinic, University of Utah Medical Center.

© 1986 by the American Society of Human Genetics. All rights reserved. 0002-9297/86/3903-0002$02.00

gene sequence, we have a direct method for defining a genetic marker located very close to a specific gene [1]: the cloned sequence serves as a probe to identify restriction fragments at the gene locus by the method of Southern [2]. The proximity of the polymorphisms (marker alleles) to the site of the hypothesized mutation within the gene, rarely more than a few thousand base pairs away, means that recombination should occur at a frequency less than 10^{-4}. Linkage analysis becomes, therefore, a precise means of defining genotypes at a specific gene locus for individuals in families.

Furthermore, the genetic linkage approach can detect mutations that affect gene expression only in a temporal or a tissue-specific fashion. These regulatory mutants might be otherwise very difficult to detect, because examination of the protein amino acid sequence or measurement of the level of the specific protein product in gross tissue might fail to reveal the mutation. However, occasionally a particular gene is implicated in the etiology of a disease and becomes the obvious choice for a linkage analysis. A probe cloned from such a candidate gene can detect regulatory mutants by linkage approaches in family studies, even under circumstances of reduced penetrance.

In diseases that are characterized by altered levels of certain proteins, linkage studies with a candidate gene could determine whether the protein alteration is a cause, or only an effect, of the disease. Moreover, if a marker with several alleles can be developed at the gene locus, linkage studies need not be large in scale; a single family with perhaps a half-dozen affected individuals may be sufficient to establish the null hypothesis if even a few recombinants are observed. Limiting such a study to a single family should reduce concerns over heterogeneity.

The group of disorders that includes coronary heart disease and hypertension is an interesting prospect for application of the candidate gene approach. There is mounting evidence that genetic predisposition is an important risk factor in these disorders, although heterogeneity and multifactorial etiology are likely [3]. One of the clearest examples of inherited predisposition to coronary heart disease is familial hypercholesterolemia (FH), which is characterized by high levels of total cholesterol and low-density lipoprotein (LDL) cholesterol; studies in many families have indicated that this disorder can result from a genetic defect in the LDL receptor protein [4, 5]. Individuals heterozygous for this receptor defect are at increased risk for early coronary disease and show a characteristic phenotype of elevated total and LDL cholesterol from early childhood, and xanthomas later in life. In some hypercholesterolemic families, xanthomas are infrequent or absent; in these cases, physicians may not be alerted to an early diagnosis. But if the elevated LDL cholesterol levels in such families are due to an inherited deficiency in the LDL receptor gene, the high levels should segregate with a specific allele of the gene, and FH could be confirmed by genetic testing at a preclinical stage, even as early as birth. We examined this hypothesis in a large Utah kindred showing a high incidence of elevated LDL cholesterol in the absence of xanthomas, using the recently cloned gene for the LDL receptor [6] to test affected individuals for linkage of elevated LDL cholesterol with the LDL receptor locus.

MATERIALS AND METHODS

Family Ascertainment

Coronary heart disease pedigree #26 (CHD 26) was among several Utah pedigrees previously ascertained through men with early (before age 55) myocardial infarction, as part of an ongoing study of coronary heart disease. This particular pedigree was selected for linkage studies because support for the segregation of a major gene for hypercholesterolemia had been found earlier by pedigree analysis [7].

Lipid Determinations

Serum total cholesterol, triglycerides, and high-density lipoprotein (HDL) cholesterol measurements were performed by the clinical laboratories at the University of Utah, using well-established techniques [8]. LDL cholesterol values were subsequently calculated from these measurements, according to the Friedewald equation [9].

Genotype Determinations at the LDL Receptor Locus

A cDNA clone of the LDL receptor gene, pHH1 [6], reveals a two-allele site polymorphism with the restriction enzyme *Pvu*II [10, 11]. Allele 1 is 16.3 kilobases (kb) long and has a frequency in unrelated individuals of .8. Allele 2 has fragments of 14.0 kb and 2.3 kb, with a frequency of .2. Individual human DNA was isolated from the nuclei of peripheral white blood cells. The genotypes were determined by hybridization of the radiolabeled, nick-translated LDL receptor cDNA probe to *Pvu*II-cut human DNA, after electrophoresis of the DNA in agarose gels and transfer to nylon membranes [12].

Linkage Analysis

Lod scores were computed using the Pedigree Analysis Package, PAP [13]. LDL cholesterol levels were adjusted for age and sex effects by regression analysis after logarithmic transformation. We assumed a mixed model of inheritance [14, 15], in which a major locus and independent polygenes determine the level of a quantitative trait. The parameters of the model, fixed at the maximum likelihood estimates obtained previously [7], were: hypercholesterolemia allele frequency of .0052; means of 124 and 304 mg/dl in the original scale for normal homozygotes and heterozygotes, respectively, and polygenic heritability of 68%.

RESULTS

Mendelian inheritance of polymorphism at the LDL receptor locus was confirmed by segregation studies in our large panel of complete, 3-generation linkage pedigrees [16]. No exceptions to Mendelian inheritance were found when 10 of these families were genotyped at the LDL receptor locus. Figure 1 shows segregation of these alleles in a selected family of CHD 26. The informative part of the CHD 26 pedigree is shown in figure 2, which indicates the individual genotypes at the LDL receptor locus and adjusted LDL cholesterol values. The distribution of age- and sex-adjusted LDL cholesterol levels in figure 2 suggests a bimodality consistent with a significant genetic component. The mean of LDL cholesterol levels for affected heterozygous individuals in CHD 26 is 328 $\pm$ 6.32 (SE) mg/dl, and for unaffected persons 145 $\pm$ 0.96 mg/dl. The corresponding levels for total cholesterol are 360 4.50 mg/dl and 209 $\pm$ 0.81 mg/dl. Triglycerides were not elevated in affected individuals, and HDL cholesterol levels were consistently lower in affected family members than in their normal counterparts.

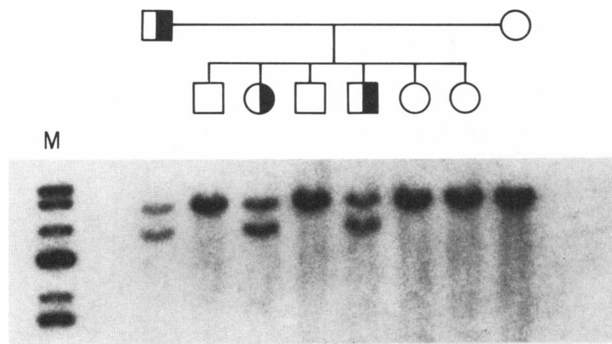

FIG. 1.—Autoradiograph of DNA sequence polymorphism showing Mendelian inheritance in a portion of CHD 26. *Half-filled symbols* represent individuals who are heterozygous for the defective LDL receptor gene and who have elevated LDL values. *The marker lane on the left* indicates fragment lengths of 18.4 kb (*top*) and 8.5 kb (*bottom*). The figure does not include the lower part of the autoradiograph that contains the constant 3.3-kb band and the smaller (2.3-kb) fragment of allele 2.

Linkage analysis demonstrated that high LDL cholesterol levels were linked at the LDL receptor locus, with a maximum lod score of 7.52 at $\theta = 0$ (fig. 3). This strongly suggested that the LDL receptor locus is, in fact, responsible for the high LDL cholesterol levels. Similar analysis using total cholesterol produced a maximum lod score of 7.99 for $\theta = 0$.

DISCUSSION

Our data provide strong support for the hypothesis that a mutant LDL receptor gene accounts for the occurrence of high LDL cholesterol levels among individuals in CHD 26. As shown in figure 3, the maximum likelihood value for θ is zero, with a steep rate of decline in likelihood with increasing recombination fraction: the chance is greater than 95% that the true value for the recombination fraction is within the bounds defined by a one-unit drop in the lod score [17]. The one-unit lod score range indicates that the high cholesterol phenotype is within a .08 recombination distance from the LDL receptor gene. The likelihood that this result would be obtained by chance between genes located on different chromosomes is less than 1%.

Another hypothesis, based on the chance that the result obtained may be due to a different gene in the vicinity of the LDL receptor, is somewhat harder to quantify but is related to the probability that the second gene would happen to be physically located within a distance that would give us the observed recombination data. If we use the 95% confidence level to define the segment, we need only calculate the probability that a gene locus would fall within a .16 recombination distance window in the genome. Taking the total genetic length of the genome to be 3,300 centimorgans, and making a few simplifying assumptions, we calculate that the probability of such a chance occurrence is less than .005.

Therefore, by either hypothesis of chance, the likelihood of observing our dataset is quite small and support for the hypothesis that the high cholesterol values are due to a defect in the LDL receptor gene is strong. Ultimately, the

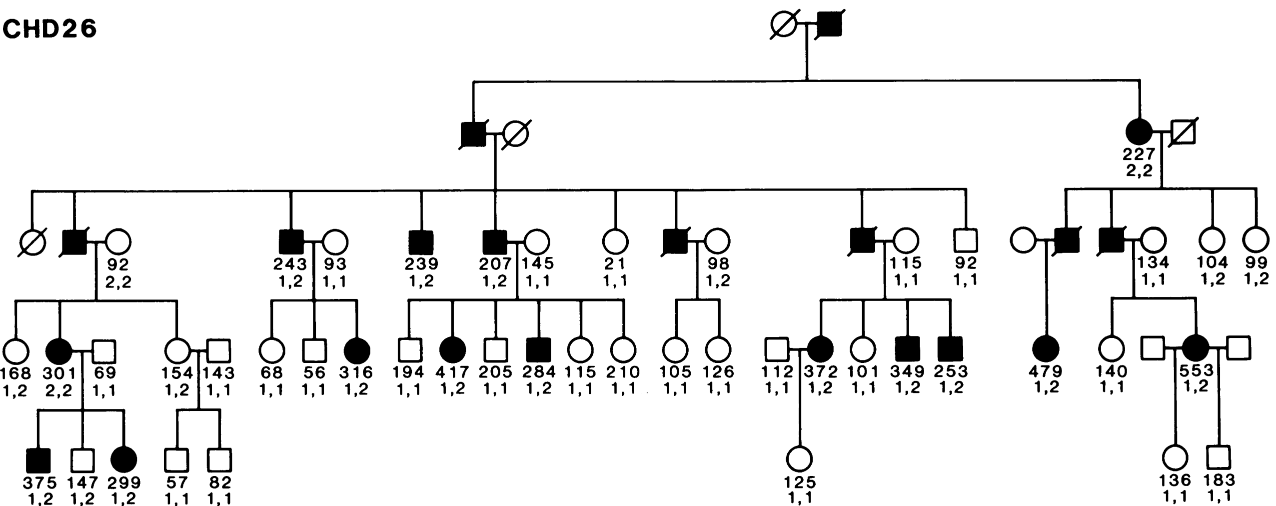

FIG. 2.—Abbreviated CHD 26 pedigree, showing only the branches with hypercholesterolemic individuals. A complete CHD 26 pedigree giving ages as well as unadjusted total cholesterol values has been published previously [7]. Lipid and genotypic measurements were also obtained on 61 unaffected relatives not included in the figure, but those data were incorporated in the linkage analysis. *The first no. under each symbol* represents the adjusted LDL cholesterol value for that individual; *the second set of nos.* represents the LDL receptor RFLP genotype. Although lod scores were computed using only the quantitative cholesterol values, affected status is designated on the pedigree as *filled circles* and *squares* when the probability of heterozygosity was > .9, based on total cholesterol levels, position in the pedigree, and RFLP genotype. *Open symbols* represent a probability of heterozygosity of < .1.

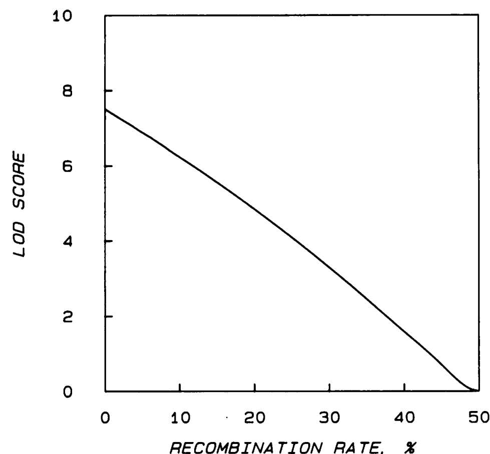

FIG. 3.—Lod score vs. recombination. Lod values for linkage between adjusted LDL levels in 107 members of CHD 26 and the LDL gene sequence polymorphism (allele 2) were calculated as described in MATERIALS AND METHODS.

issue of chance association vs. mutation may be resolved by careful determinations of LDL receptor activity in fibroblasts or by the demonstration that the cloned LDL receptor gene from this family has a mutational lesion adequate to account for the phenotype. The tight linkage we have observed rules out many other genes known to be involved in lipid metabolism as major contributors to hypercholesterolemia in CHD 26, because these other candidates do not map close to the LDL receptor gene on chromosome 19p [17]. Occasionally, however, phenotypically similar families might be found in which the biochemical etiology for FH is different, and such families would be expected to show no linkage to the LDL receptor.

It is noteworthy that the entire study was accomplished within a single kindred, since the possibility of heterogeneity was markedly reduced with such a strategy. Furthermore, by doing the study within a single kindred, we maintained knowledge of the allelic phase relationships, considerably increasing the efficiency of the study. However, Humphries et al. [11] were able to establish tight linkage between the LDL receptor gene and hypercholestrolemia by combining linkage data from two families; they obtained a lod score of 3.6 at $\theta = 0$. The defective LDL receptor gene cosegregated with the more frequent 16.3-kb allele in both families of that study, while in CHD 26, the cosegregating marker was the less frequent allele.

The single-site polymorphism at the LDL receptor locus that was used for the linkage studies in CHD 26 clearly is inadequate at the present time for routine clinical or research purposes in many small FH families because the allele frequencies are .8 and .2. However, development of other useful polymorphisms awaits only the discovery of RFLPs in sequences adjacent to the pHH1 clone.

CHD 26 was chosen for our initial study because its structure and phenotypic pattern were optimal for the purpose. Nevertheless, we anticipate that the approach is sufficiently powerful to reveal linkage even when the bimodal

distribution of cholesterol values is not as obvious. In particular, it would be instructive to examine large families that show only moderately increased cholesterol levels in order to investigate the possible etiologic involvement of mutant LDL receptor alleles showing milder phenotypic expression. In any event, the ability to unambiguously determine genotype at a defective LDL receptor locus will help investigators to elucidate the effects of environmental factors such as drug therapy or diet on phenotypic expression of the mutant gene, as well as possible interactions between the LDL receptor locus and other genes. Variability in expression of disease in persons with the mutant genotype is suggested by the observation that four affected male ancestors of extended CHD 26 lived to ages 62–81 without early coronary disease [7].

The findings reported here validate the candidate gene approach to the study of coronary heart disease; moreover, they are likely to have significant usefulness to members of CHD 26. Presumably, life-style changes and genetic counseling could lessen the impact of the disease on the carriers of the defective gene in this family and on their progeny. On the other hand, when individuals, of whom there were five in CHD 26, are distinguished as a normal genotype with the DNA marker but exhibit marginally elevated cholesterol levels, counseling by the clinician need not take into account the risk of passing FH to offspring.

ACKNOWLEDGMENTS

We gratefully acknowledge the kind gift of the LDL receptor clone from Drs. Helen Hobbs, Michael Brown, and J. L. Goldstein. We also wish to thank Mary B. Christensen, Dorothy Hudman, and Barbara Ogden for their help in collecting blood samples; Mary Carlson, Mary Dadone, Tami Elsner, Mary Hadley, and Lesa Nelson for expert technical assistance; and Ruth Foltz for editorial assistance in the preparation of this manuscript.

REFERENCES

1. Kan YW, Dozy AM: *Proc Natl Acad Sci USA* 75:5631–5635, 1978
2. Southern E: *J Mol Biol* 98:503–517, 1975
3. Williams R: In *Update IV: The Heart,* edited by Hurst J, New York, McGraw-Hill, 1979, pp 89–118
4. Brown M, Goldstein JL: *Science* 185:61–63, 1974
5. Goldstein JL, Dana SD, Brunschede GY, Brown MS: *Proc Natl Acad Sci USA* 72:1092–1096, 1975
6. Yamamoto T, Davis CG, Brown MS, et al.: *Cell* 39:27–38, 1984
7. Williams RR, Hasstedt SJ, Wilson MD, et al.: *J Am Med Assoc* 255(2):219–224, 1986
8. Rautela GS, Liedtke RJ: *Clin Chem* 24:108–114, 1978
9. Friedewald WT, Levy RI, Fredrickson DS: *Clin Chem* 18:499–502, 1972
10. Hobbs H, Lehrman MA, Yamamoto T, Russell DW: *Proc Natl Acad Sci USA* 82:7651–7655, 1985
11. Humphries SE, Horsthemke B, Seed M, et al.: *Lancet* i:1003–1005, 1985
12. Barker D, Schafer M, White R: *Cell* 36:131–138, 1984
13. Hasstedt SJ, Cartwright PE: Technical Report No. 13, Dept. of Medical Biophysics and Computing, Salt Lake City, Univ. of Utah, 1981
14. Elston RC, Stewart J: *Hum Hered* 21:523–542, 1971
15. Morton NE, MacLean CJ: *Clin Chem* 24:108–114, 1974
16. White RL, Leppert M, Bishop DT, et al.: *Nature* 313:101–105, 1985
17. Conneally P, Edwards J, Kidd K, et al.: *Cytogenet Cell Genet* 40:356–359, 1985