Four DNA polymorphisms in the LDL receptor gene: their genetic relationship and use in the study of variation at the LDL receptor locus

ROHAN TAYLOR*, MOHAMED JEENAH*, MARY SEED†, AND STEVE HUMPHRIES*

From *The Charing Cross Sunley Research Centre, Lurgan Avenue, London W6 8LW; and †Charing Cross Hospital Medical School, St Dunstan’s Road, London W6 8RF.

SUMMARY We have studied four different restriction fragment length polymorphisms (RFLPs) for the LDL receptor gene, detected using the restriction enzymes Stul, PvulI, ApaLI, and NcoI, in normal subjects and in patients with familial hypercholesterolaemia (FH) from London. Significant linkage disequilibrium was detected between all four RFLPs. Used together they give a polymorphism information content (PIC) of >0.7 which makes them useful for studying the inheritance of the LDL receptor gene in more than 70% of families with FH. The NcoI and ApaLI RFLPs were found to be the most useful, giving a combined PIC value of 0.6. The allele frequencies of all four polymorphisms were compared in the normal and FH groups and the frequency of the rarer N2 allele of the NcoI RFLP was found to be significantly higher in the FH group. This suggests that a mutation has occurred on the rare NcoI N2 allele and that it may be making a significant contribution to the defects causing FH in this patient group. We have also used these RFLPs to look for evidence that variation at the LDL receptor gene locus contributes to the determination of cholesterol levels in the normal population. People with different RFLP genotypes do not have significantly different levels of serum total or LDL cholesterol. At present we have no evidence that variation at this locus may be determining cholesterol levels in the non-FH population.

Familial hypercholesterolaemia (FH) is a common autosomal dominant disease with a heterozygote frequency of about 1 in 500. It is clinically characterised by raised serum and LDL cholesterol levels, tendon xanthomata, and an increased risk of myocardial infarction after the second decade of life.1-3 Over recent years it has been shown that FH is caused by mutations in the gene for the low density lipoprotein (LDL) receptor.4 5 This receptor is found on the surface of all cells and its normal function is to bind and internalise the cholesterol rich LDL particles. Defects in receptor function lead directly to the raised blood cholesterol levels which cause the development of atherosclerosis and premature coronary artery disease. It is therefore highly advantageous to identify affected subjects at a young age before they develop symptomatic disease, in order to give advice on diet and prophylactic treatment with drugs to reduce their future risk of myocardial infarction. It has been shown that measurement of total or LDL cholesterol levels does not always allow unequivocal diagnosis of FH, especially in children.6 7 since these values may fall in the upper range of normal. One study6 estimated that 4-25% of children with one parent heterozygous for FH could not be diagnosed equivocally by these means alone.

The cloning of the LDL receptor gene8 has now made it possible to analyse mutations causing FH at the DNA level. To date, at least 16 different mutations have been described in detail,9-19 but there are likely to be many more than this number. Gross structural alterations to the gene, such as deletions or insertions, can be detected directly using Southern blotting techniques and we have recently shown that about 4% of FH patients in London have these detectable deletions.19 However, the majority of defects are probably the result of point mutations and will require other methods for diagnosis.
Over the last two years, 10 restriction fragment length polymorphisms (RFLPs) in the LDL receptor gene have been reported using eight different enzymes.20-27 These can be used as markers to follow the inheritance of a defective receptor gene indirectly in a family by linkage analysis.27-29 Thus, in families where some members have FH, it may be possible to carry out unequivocal diagnosis in children whose age adjusted serum cholesterol levels are within the normal range, but who have inherited a defective LDL receptor gene. In this study we have examined the usefulness of four common RFLPs of the LDL receptor gene, all detected using the same probe, for the diagnosis of FH in the London population.

It has recently been shown that variation at the apolipoprotein (apo) E locus contributes as much as 8% of the phenotypic variance in cholesterol levels in the general population.30 31 Therefore, it is likely that variation in other genes that are involved in lipid metabolism, such as apo B or the LDL receptor, may also contribute to the determination of cholesterol levels in the normal population. We have attempted to study the effect of variation at the LDL receptor locus on normal lipid levels by examining the mean cholesterol levels of subjects with different LDL receptor RFLP genotypes.

Materials and methods

Blood samples

Blood samples were collected from normal healthy Caucasian men and women representative of the non-FH population in London, who had completed a Rose questionnaire to exclude the presence of ischaemic heart disease. They were aged 45 to 59 years with a mean total cholesterol of 5-6 (SD 1-02) mmol/l and mean triglyceride 1-7 (SD 0-95) mmol/l. This normal population included subjects with sporadic hypercholesterolaemia but excluded persons with familial hypercholesterolaemia. The other characteristics of this control sample will be presented elsewhere (R Houlston et al, in preparation). Blood samples were also collected from Caucasian patients with heterozygous FH, attending the lipid clinics at Charing Cross and Hammersmith Hospitals. Standard criteria for diagnosis of FH were used.32 Briefly these were: total serum cholesterol levels above 6-7 mmol/l owing to serum LDL cholesterol levels above 4-9 mmol/l, accompanied by tendon xanthomata and xanthomata or hypercholesterolaemia in a first degree relative. A total of 10 ml of whole venous blood was collected into EDTA tubes and stored at -20°C until DNA extraction was carried out.

Probes for the LDL receptor gene

The 3' gene probe used was pLDLR-2HH1, a kind gift from Dr D W Russell, Dallas, USA. This plasmid contains a 1-9 kb fragment of the full length receptor cDNA (bp 1573 to 3486). This is equivalent to the designated 1-9 kb BamHI fragment of the full length receptor cDNA pLDLR3. The insert was cut out by digestion with BamHI (Anglian Biotech) and separated from the vector fragment by electrophoresis on a 1% low melting point agarose gel. The 1-9 kb fragment was excised from the gel and boiled in water. The probe was labelled to a specific radioactivity of >10⁸ cpm/μg by random hexanucleotide priming using [³²P] dCTP at 800 Ci/mmol (Amersham).

DNA analysis

Total genomic DNA was extracted from white blood cells using a Trition X-100 lysis method. A total of 5 μg of DNA was digested using the restriction enzymes Stul, PvuII, ApaLI plus BamHI, and NcoI (Anglian Biotech) at a concentration of 2 to 4 U/μg DNA. Restriction fragments were size separated by electrophoresis on 0-8% or 1% agarose gels and transferred to Hybond-N nylon membranes (Amersham) by Southern blotting.34 DNA was fixed by UV treatment or baking at 80°C and filters were prehybridised for at least 10 minutes at 65°C in 7% sodium dodecyl sulphate (SDS), 0-5 mol/l phosphate buffer, pH 7-0. Hybridisation was carried out in 50 ml tubes in rotating ovens (Bachofer, West Germany) at 65°C in 3 ml of the same solution plus 3×10⁵ cpm of denatured radiolabelled probe. After at least 16 hours' hybridisation, the filters were washed to a stringency of 1×SSC, 0-1% SDS for 30 minutes at 65°C. Filters were exposed to x ray film using an intensifying screen at -70°C for one to three days.

Calculation of Δ

The correlation coefficient Δ was used to determine the degree of linkage disequilibrium between two RFLP loci A and B. It was calculated from the equation below:

\[
\Delta = \frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}
\]

a, b, c, and d being the frequencies of the haplotypes A1B1, A1B2, A2B1, and A2B2 respectively. The 95% confidence limits of Δ were calculated using the Z transformation where Z=tanh⁻¹ Δ and Z has normal distribution with standard error of

\[
\frac{1-96}{\sqrt{n-3}}
\]
Pairwise PIC values were calculated from the observed frequency of the haplotypes in the control population. For subjects heterozygous for both polymorphisms, the relative frequency of the two possible haplotypes was determined from the frequency of the unambiguously observed haplotypes, using an estimation maximisation algorithm.\textsuperscript{35}

**Statistical Analysis**

χ² analysis was used to test the hypothesis of random association between the alleles of two particular RFLPs. Statistical significance was considered to be at the 0.05 level. The lipid variables were corrected for height, weight, age, and sex differences by polynomial regression. One way analysis of variance was used to test for association of the RFLPs with phenotypic variation. Analysis was carried out using the MINITAB program (State College, Pennsylvania).

**Results**

The LDL receptor gene fragments detected by the 3' receptor cDNA probe in Southern blot hybridisations using the four different enzyme digests are shown in fig 1. Fig 2 shows a map of the 3' half of the LDL receptor gene with positions of the varying enzyme sites marked and the approximate sizes of the hybridising fragments shown below the map. The varying StuI site is the nearest to the 5' end of the gene, while the ApaLI and PvuII varying sites are close together in intron 15. Our preliminary mapping data indicate that the varying NcoI site is at the 3' end of the gene.

Table 1 shows the genotype distribution and relative allele frequency in both the normal and FH sample examined for all four RFLPs. None of the genotype distributions differed significantly from that expected for a sample in Hardy Weinberg equilibrium. The frequency of the PvuII, StuI, and ApaLI RFLPs did not differ significantly between the two populations. However, the frequency of the common allele (N1) of the NcoI RFLP is 0.74 in the control sample but a lower frequency of 0.63 is observed in the sample of FH subjects, and this difference is statistically significant (χ²=5.20, 1 df, p<0.025 by gene counting).

In table 2 correlation coefficient (Δ) values are presented for all six pairwise combinations of the four polymorphisms with the confidence limits to examine whether there is a significant deviation from Δ=0 (that is, linkage disequilibrium). The Δ values are similar in the control group and in the patients with FH (not shown). The pairwise polymorphism information content (PIC) values are also presented in table 2. The highest value is 0.61 for the NcoI and ApaLI polymorphisms, which makes these RFLPs the most informative when used together. The StuI polymorphism has the lowest rare allele frequency and a PIC value of 0.1 (table 1) when

**Fig 1** Southern blot hybridisation with the 3' LDL receptor cDNA probe p2HH1 using four different restriction enzyme digests, StuI, PvuII, ApaLI/BamHI, and NcoI, to detect the four RFLPs. In each we have designated the rare allele S2, V2, A2, and N2.
TABLE 1  Genotype distribution and relative allele frequency of the Stul, PvuII, ApaLI, and NcoI RFLPs in controls and FH patients from London.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Relative allele frequency</th>
<th>PIC value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stul controls (n=154)</td>
<td>S1/S1 138, S1/S2 145, S2/S2 145</td>
<td>0-10</td>
</tr>
<tr>
<td>FH (n=61)</td>
<td>S1 56, S2 95, S/S 0-940-06</td>
<td>0-90-04</td>
</tr>
<tr>
<td>PvuII controls (n=147)</td>
<td>V1/V1 79, V1/V2 95, V2/V2 95</td>
<td>0-31</td>
</tr>
<tr>
<td>FH (n=59)</td>
<td>V1 33, V2 67, V/V 0-750-25</td>
<td>0-760-24</td>
</tr>
<tr>
<td>ApaLI controls (n=66)</td>
<td>A1A1 18, A1A2 33, A2A2 15</td>
<td>0-520-48</td>
</tr>
<tr>
<td>FH (n=53)</td>
<td>A1 16, A2 87, A/A 0-560-44</td>
<td>0-560-44</td>
</tr>
<tr>
<td>NcoI controls (n=117)</td>
<td>N1N1 66, N1N2 33, N2N2 15</td>
<td>0-31</td>
</tr>
<tr>
<td>FH (n=126)</td>
<td>N1 50, N2 61, N/N 0-630-37*</td>
<td>0-630-37*</td>
</tr>
</tbody>
</table>

*χ²=5-20, 1 df by gene counting (p<0-025).

used on its own, and it does not provide significant extra information when combined with any of the other RFLPs.

Since PIC values are highest for the ApaLI and NcoI RFLPs (table 1), we have compared the relative genotype frequency when using the four RFLPs in conjunction with using the NcoI and ApaLI RFLPs only (table 3). When using the two RFLPs only, 71% of the control population and 84% of the FH group are heterozygous for one or both polymorphisms. Using the four RFLP genotypes, we found that 76% of the controls and 86% of the FH group were heterozygous for at least one

TABLE 2  A values and PIC values of the pairwise combinations of four RFLPs.

<table>
<thead>
<tr>
<th>Δ value*</th>
<th>95% confidence limits</th>
<th>No</th>
<th>Combined PIC value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PvuII-NcoI</td>
<td>0-35</td>
<td>69</td>
<td>0-57</td>
</tr>
<tr>
<td>PvuII-Stul</td>
<td>0-43</td>
<td>69</td>
<td>0-37</td>
</tr>
<tr>
<td>PvuII-ApaLI</td>
<td>0-63</td>
<td>66</td>
<td>0-54</td>
</tr>
<tr>
<td>NcoI-Stul</td>
<td>0-18</td>
<td>69</td>
<td>0-43</td>
</tr>
<tr>
<td>NcoI-ApaLI</td>
<td>0-35</td>
<td>66</td>
<td>0-61</td>
</tr>
<tr>
<td>Stul-ApaLI</td>
<td>0-30</td>
<td>66</td>
<td>0-47</td>
</tr>
</tbody>
</table>

*All Δ values are significantly different from 0 (p<0-05), that is, linkage disequilibrium.

FIG 2  Map of the 3' half of the LDL receptor gene showing the positions of the varying sites of the Stul (S), PvuII (V), ApaLI (A), and NcoI (N) RFLPs. The fragments that are detected by probe p2HHI are shown below. BamHI sites are marked B. The cDNA probe p2HHI spans the exons from the BamHI site in exon 10 to the BamHI site in exon 18. We used double NcoI/Stul digests to map the varying NcoI site to the Alu repeat region of exon 18. The 13 kb NcoI fragment is reduced to 3-6 kb by the Stul cutting site in exon 18 and the 9 kb NcoI fragment is reduced to 5 kb by the Stul site in intron 15. The 7 kb and 3-4 kb NcoI fragments remain uncut by Stul digestion. This map has been confirmed by the use of an exon 13 and 14 specific cDNA probe (cut out of p2HHI) which hybridises to the 7 kb NcoI fragment only (not shown).
polyorphism and therefore potentially informative for diagnosis of FH.

We have also used the polymorphisms to see if we can detect significant differences in cholesterol levels in groups of subjects with different RFLP genotypes. Table 4 gives the mean cholesterol levels in normal subjects with different LDL receptor RFLP genotypes. When one way analysis of variance was carried out, no significant differences in cholesterol levels were found associated with the different genotypes, although subjects with the S2 allele have lower levels of cholesterol. No differences in LDL cholesterol or apo B levels were found in subjects with different RFLP genotypes (data not shown).

### Discussion

There are now several effective lipid lowering drugs and therapies available to treat patients with FH and to reduce their risk of coronary artery disease. Once detected either by conventional biochemical methods or by genetic tests, affected subjects can be given advice with regard to other risk factors (smoking, diet, and exercise) and prophylactic drug therapy for the hypercholesterolaemia to lessen the risk of development of premature atherosclerosis. One of the aims of this study was therefore to determine the usefulness of four RFLPs of the LDL receptor gene as linkage markers for diagnosis in families with known FH. The usefulness of such polymorphisms for family (linkage) studies depends on the relative frequency of the alleles and the degree of linkage disequilibrium between the alleles of the different polymorphisms.

The frequency of the PvuII RFLP in this sample is the same as previously reported for the London population. The StuI RFLP was first reported in the South African Afrikaner population where the frequency of the common S1 allele was found to be 0.92, which is not significantly different from the frequency we observe here. The NcoI RFLP was also reported in South African Caucasians where the frequency of the N1 allele is 0.64 compared to 0.74 in the London population. This difference is not statistically significant and might be the result of chance variation in the small samples studied, or may reflect the different ethnic origins of the two populations. The ApaLI RFLP in intron 15 was reported to have a frequency for the 6-6 kb allele (A1) of 0.43 when measured in 13 American Caucasians, but we have found the frequency of the A1 allele in the London sample to be 0.52, which makes it more useful than previously thought.

The four polymorphisms are all clustered in the 3' half of the gene with the StuI site being nearest to the 5' end in exon 8 and the NcoI site nearest to the 3' end in exon 18. If there were linkage equilibrium (random association of the alleles) between the four
RFLPs, we would expect their combined PIC value to be greater than 0.9. However, there is significant linkage disequilibrium between all of them which reduces the overall PIC value to approximately 0.7. Over such small genetic distances the degree of linkage disequilibrium is not related primarily to the physical distance between the varying genetic markers, but is rather the result of the evolutionary history of the four polymorphisms. However, the data are consistent with the physical map of the location of the polymorphisms.

For linkage studies, most of the information can be obtained by using only the two polymorphisms ApaLI and NcoI, whose combined PIC value is 0.61. This means that, using these two RFLPs alone, roughly two thirds of FH families will be informative for genetic studies. It is likely that some of the other RFLPs that have been recently reported, but not used in this study, will be additionally useful for early diagnosis of FH in families or occasionally for prenatal diagnosis where both parents have heterozygous FH.

We have also compared the frequency of LDL receptor RFLPs in FH patients and the normal population in London. We have used four RFLPs and have observed a statistically significant difference only in the allele frequency of the NcoI RFLP. The absence of the varying cutting site (rarer N2 allele) occurs more frequently in the FH group than in the controls. This suggests that in our London FH sample, a mutation has occurred on an N2 allele and is making a significant contribution to mutations causing FH in this patient group. From this we would predict that a large proportion of the patients possessing a defective LDL receptor gene associated with an N2 allele will have a common mutation, and they may therefore share a common haplotype using the other RFLPs. Further haplotype analysis and receptor binding studies are being carried out to confirm this prediction. It is, however, possible that this observation could have occurred by chance alone and this will need to be confirmed in larger studies. It is possible that the frequency of these four RFLPs is different in populations from different geographical areas. This is the case for polymorphisms of the apo AI/CIII/AIV gene cluster, which show significant differences in allele frequency even within different parts of the UK. Thus, if the origins of the patient and control groups are not the same, this could lead to the observed differences in frequency. Our patients and controls are all Caucasians from London, but they probably originate from several different European countries.

We have also attempted to use these RFLPs to see if variation in the LDL receptor gene may contribute to the determination of normal serum cholesterol levels in non-FH hyperlipidaemic and normal subjects. For example, a variation in the promoter region of the gene might result in slightly reduced transcription of the gene and thus lower receptor numbers on the cell surface. Alternatively, amino acid changes in the ligand binding region might slightly reduce the affinity for binding LDL, and both of these changes might predispose a subject to develop moderately raised levels of LDL cholesterol. Some family data have been reported to support the hypothesis that four different 'normal' LDL receptor variants exist, but this has not been confirmed. We have determined the mean cholesterol levels of persons in the normal population with different LDL receptor RFLP genotypes to study the effect of variation at this locus on normal lipid levels. This approach has proved fruitful in the analysis of variation associated with the apo B and apo AI polymorphisms. However, in our control sample there are no statistically significant differences in mean serum cholesterol (table 4), LDL cholesterol, or apo B levels associated with different StuI, PvuII, ApaLI, or NcoI genotypes. Some differences in cholesterol levels associated with the StuI RFLP were noted, and it is possible that these differences would reach statistical significance if repeated in a larger sample. This failure to detect a significant association may be because the RFLPs used are all in the 3' region of the gene and it is possible that other polymorphisms, particularly those nearer the promoter and ligand binding regions at the 5' end of the gene, would be more useful. At present we have no evidence that variation at the LDL receptor locus is involved in determining cholesterol levels in the non-FH population.

This work was supported by grants from the British Heart Foundation (RG5), the Family Heart Association, and The Charing Cross Sunley Research Trust. MSJ is supported by a grant from the South African MRC. We would like to thank Dr D W Russell for supplying the LDL receptor probe, Dr Gil Thompson for access to patient samples, and Professor M Marmot, Dr R Houlston, and Dr D Wile for supplying DNA samples. We are also grateful to Miss Naiia Loqueman for skilled technical assistance and to Mrs A Wilcox for help in preparing the manuscript.

References


LDL receptor gene polymorphisms


Lehrman MA, Goldstein JL, Brown MS, Russell DW, Schneider WJ. Internalization-defective LDL receptors produced by genes with nonsense and frameshift mutations that truncate the cytoplasmic domain. Cell 1985;41:735–43.


Templeton AR, Sing CF, Kessler A, Humphries SE. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. II. Associations of triglyceride and cholesterol levels with the apolipoprotein AI, III, AIV region in a human population. Genetics (in press).


Correspondence and requests for reprints to Miss R Taylor, The Charing Cross Sunley Research Centre, Lurgan Avenue, London W6 8LW.
Four DNA polymorphisms in the LDL receptor gene: their genetic relationship and use in the study of variation at the LDL receptor locus.

R Taylor, M Jeenah, M Seed and S Humphries

doi: 10.1136/jmg.25.10.653

Updated information and services can be found at:
http://jmg.bmj.com/content/25/10/653

**Email alerting service**

*These include:*

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/