Identification of a common low density lipoprotein receptor mutation (R329X) in the south of England: complete linkage disequilibrium with an allele of microsatellite D19S394


Abstract
Familial hypercholesterolaemia is commonly caused by mutations in the low density lipoprotein receptor (LDLR) gene and more than 300 different mutations have been described worldwide. Some mutations occur at relatively higher frequency in certain populations, reflecting both chance and demography, most evident in founder populations. As part of a study of kindreds of 78 probands from Southampton and south west Hampshire, we identified the same mutation (R329X) in 978 (11.5%) probands. In all (100%) of these probands, length allele 259nt of the 17 allele microsatellite D19S394, sited approximately 250 kilobases telomeric and 5' to the LDLR gene, was observed, although in the general population this allele has a prevalence of only 16.1%. A simple diagnostic assay for R329X was constructed in conjunction with more detailed family studies. Both the R329X and linked D19S394 allele also cosegregated with the FH phenotype within each kindred. Although R329X involves a CpG site, it is highly likely that the families are identical by descent for R329X, we surmise with a common ancestor within 500 to 1000 years, although the mutation is not restricted to this geographical area. This relationship illustrates that the linkage disequilibrium of gene LDLR with marker D19S394 will enable rapid recognition using D19S394 genotype of possible common FH mutation(s) within a cohort of FH patients from a particular locality or ethnic group.

Keywords: LDL receptor; familial hypercholesterolaemia.

Familial hypercholesterolaemia (FH) is a common inherited disease showing an autosomal dominant pattern of inheritance defined as being attributable to molecular defects in the LDL receptor. It is characterised clinically by a rise in the concentration of low density lipoprotein (LDL) cholesterol in blood, tendon xanthomata, and an increased risk of myocardial infarction. Based on the estimated population frequency of carriers of 1/500, there are more than 110 000 FH heterozygotes in the UK, of whom probably fewer than 30-50% are even known to have hypercholesterolaemia, and fewer than 0.1% have been established at the genetic level to carry a gene causing familial hypercholesterolaemia. The hyperlipidaemia and arterial lesions of these patients are responsive to treatment by diet and drugs, and such treatment is likely to reduce subsequent morbidity and mortality. Although management is based on the phenotype, counselling and future family tracing programmes will be helped by knowledge of the genotype so that definitive genetic tests can be offered. Genetic characterisation would thus be helpful both for research and in the longer term as a diagnostic tool. Given a more complete knowledge of the spectrum of mutations, it may be possible to assemble cost effective direct assays for FH.

Worldwide, more than 300 different mutations of the LDLR gene have been characterised at the DNA level. Within a geographically or culturally isolated population, or where a large proportion of people are related by descent of migration, there may be a single mutation causing FH in many of the patients. In the UK, where there is a heterogeneous population, it is unlikely that any mutations will be present at a high frequency in FH patients. Findings to date in a sample of 200 FH patients from London are that 5% of patients have a gross rearrangement, while in a further 18% a small deletion or a single base mutation has been detected in exon 3, 4, or 14. However, there may be increased prevalence of some mutations in some regions. So far in England, only one such mutation has been identified, which accounts for approximately 10% of FH in the Manchester area, and which is widely distributed in the UK with a more typical prevalence of 1-2% among FH patients. However, many of the reported studies have been on patients from London clinics, which have a highly heterogeneous mix of patients. The objectives of this study were to identify any mutation which might be common in the south of England and to examine highly polymorphic microsatellite D19S394 on chromosome 19p, currently under study as a marker for cosegregation studies of FH families (L Haddad, I Day, in preparation), for linkage disequilibrium with any common mutation found.
Methods

PROBANDS AND FAMILIES

Probands fulfilled standard diagnostic criteria for FH.1 Family tracing was undertaken on a research basis according to approval from the local ethical committee.

SAMPLE ACQUISITION AND PREPARATION

DNA from probands and from some relatives was purified as described previously from standard 5 ml potassium-EDTA venous blood.10 For some relatives, DNA was purified as described previously11 from mouthwash samples obtained by postal return. Mouthwash samples were obtained in 10 ml aliquots after written consent and telephone interview. Patients were supplied with a 20 ml sterilin tube containing 10 ml of sterile isotonic saline, which they were instructed to draw into their mouth and swill for 30 seconds before spitting back into the tube. Sterilin tubes were supplied and returned in breakproof packaging which fits most letter boxes. Mouthwash samples were found to yield very good PCR template DNA even after one to three days transit at ambient temperature or storage at -20°C before DNA extraction or both.

SSCP ANALYSIS AND SEQUENCING

SSCP and sequencing were undertaken as described previously for a study of LDLR gene exon 3,13 except that primers for LDLR gene exon 7 were used in the present study.

ANALYSIS OF POLYMORPHIC MICROSATellite

D19s394

PCR oligonucleotide sequences for the tetranucleotide repeat marker D19s394 were as follows: 5'-FAM-AGACTACAGTGAGCTG-TGG for the sense primer, and 5'-GTGTTCCTAACTACCAGGC for the antisense primer (Genome DataBase, http://gdbwww.gdb.org). Primers were from Genosys, Cambridge, UK. Optimal PCR conditions were: a denaturation step for one minute at 94°C, followed by an annealing temperature of 58°C for one minute and an extension of 72°C for two minutes for the first five cycles, followed by 30 cycles with the same conditions but with an annealing temperature of 54°C. MgCl2 was at a final concentration of 1.0 mmol/l for PCR, and Tween 20 (Aldrich Chemical Co Ltd, UK) at 0.5% v/v. PCR reactions were done in a total volume of 20 μl and overlaid with 20 μl paraffin oil. Each 20 μl reaction mix contained 8 pmol of each PCR primer, 0.2 U Taq polymerase (Gibco-BRL Ltd, Paisley, UK), 200 μmol/l each dNTP, 50 mmol/l Tris-HCl (pH 8.3), 0.01% (w/v) gelatin. Thermal cycling was in a Hybaid Omnigene Thermal Cycler (Hybaid, Teddington, Middlesex, UK).

RESTRICTION SITE ASSAYS FOR R329X

BstI digests the sequence CCAGT and therefore enables a specific assay for mutation R329X, in which the normal sequence CCAGC is mutated to CCAGT. Consistent with the manufacturer's recommendations (New England Biolabs, personal communication), it was found essential to purify the PCR product from the constituents of the PCR mix before digestion. Using PCR primers as used for SSCP, normal sequence yields a 169 base pair product, but this is cleaved to 128 and 41 base pair products only if the R329X mutation is present.

The use of two primers sited directly adjacent to the CpG sequence of codon 329 (TpG for R329X), in which the 3' bases of both primers is T, leads to PCR induction of a TaqI site in the normal sequence. Any mutation at the CpG site, including R329X, will ablate the TaqI site produced, yielding a sensitive but not necessarily specific assay for R329X. This approach has been developed into a general system for analysis of any CpG site,14 combining up to two 3' mismatched PCR primers, very short PCR (that is, oligonucleotides plus intervening CpG), TaqI digestion, and polycrylamide microplate array diagonal gel electrophoresis (MADGE) analysis.15 Among other CpG mutations and forcing of other four base restriction sites containing CpG, the TaqI assay for R329X is considered in detail in reference 17.

Results

PROBANDS POSSESSING R329X MUTATION

Probands identified were a subset of those described previously and fulfilled standard criteria for probable or definite familial hypercholesterolaemia. Probands' ages ranged from 17 to 46 years and untreated cholesterol levels were from 9.5 to 12.3 mmol/l. Two probands had xanthomata, one knew of xanthomata in an uncle, two had distinctive conreal arbus, and one had distinctive xanthelasma. In every case the cholesterol level was approximately twice the reference value for age (table 1), but triglyceride levels were not raised and there was no identifiable secondary cause for the hypercholesterolaemia, and in each family there was bimodal segregation of cholesterol levels.

IDENTIFICATION OF MUTATION R329X,
CONFIRMATION IN PROBANDS, AND ANALYSIS OF RELATIVES

Single strand conformation polymorphism analysis of all exons of probands showed a variant pattern in exon 7, apparently identical in nine people (fig 1). Sequencing of several of them showed the same base change, C to T at nucleotide 1048 (fig 1). This base change results in a nonsense mutation at codon 329, converting arginine to a stop codon (R329X). A direct assay using PCR to induce a TaqI restriction site17 was used in conjunction with SSCP to test for and confirm or exclude the presence of the mutation R329X in all probands from the region and in all relatives studied from the families of established R329X probands (fig 1). Nine patients were positive for R329X out of 78 probands tested (11.5%). Extensive family tracing, including five generation histories for most families, resulted in the eventual reduction of the dataset to six large, apparently independent kindreds. For one kindred, the apparent FH ancestor originated in York, not in the south of England.
Common LDLR gene mutation

Table 1  Baseline lipid measures for FH probands for R329X families and molecular data for probands and relatives

<table>
<thead>
<tr>
<th>Proband</th>
<th>Unretrated TC (mmol/l), age (s), gender (M/F)</th>
<th>HDL-Chol (mmol/l)</th>
<th>Fasting TG (mmol/l)</th>
<th>D19S394 allele sizes (nt)</th>
<th>No of relatives revised</th>
<th>No of relatives positive for FH by TC</th>
<th>No of relatives proven R329X positive</th>
<th>No of relatives with R329X + D19S394 259nt allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>10.2, 29, M</td>
<td>0.78</td>
<td>2.2</td>
<td>+</td>
<td>247, 259</td>
<td>9</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>010</td>
<td>10.5, 44, F</td>
<td>0.77</td>
<td>0.8</td>
<td>+</td>
<td>227, 259</td>
<td>11</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>012</td>
<td>10.1, 31, M</td>
<td>N/AV</td>
<td>1.2</td>
<td>+</td>
<td>259, 263</td>
<td>14</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>074</td>
<td>12.3, 45, M</td>
<td>0.79</td>
<td>2.1</td>
<td>+</td>
<td>255, 259</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>016</td>
<td>10.3, 46, F</td>
<td>1.06</td>
<td>1.4</td>
<td>+</td>
<td>231, 259</td>
<td>7</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>061</td>
<td>9.9, 31, M</td>
<td>N/AV</td>
<td>1.1</td>
<td>+</td>
<td>259, 267</td>
<td>16</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>045</td>
<td>11.1, 17, F</td>
<td>1.13</td>
<td>0.8</td>
<td>+</td>
<td>227, 259</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>055</td>
<td>10.1, 22, F</td>
<td>1.26</td>
<td>0.8</td>
<td>+</td>
<td>251, 259</td>
<td>8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>074b</td>
<td>9.5, 24, F</td>
<td>0.94</td>
<td>0.6</td>
<td>+</td>
<td>259, 295</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

N/AV = not available. *074b was an index case in this study, but turned out to be a first degree relative of 074 (different name and minimal contact). *074b was established to be genotype 259, 295nt for D19S394 (not 259, null) from family study: direct genotyping showed a single peak at 259nt.

Identification of linkage disequilibrium between polymorphic microsatellite D19S394 and R329X

Simultaneously with our work examining the spectrum of LDLR gene mutations in familial hypercholesterolemics from the south of England, we have been examining the use of polymorphic microsatellite loci flanking the LDLR gene for rapid cosegregation studies of apparent FH kindreds (L Haddad, 1 Day, in preparation). During this study it was noted that not only did a D19S394 allele cosegregate with the FH phenotype in each R329X family (fig 2), but also that there was complete linkage.

---

Figure 1  Molecular characterisation of LDLR gene mutation R329X. (A) SSCP analysis of LDLR gene exon 7 PCR product from normal and R329X heterozygotes. The latter display additional bands representing the R329X allele PCR product. (B) Sequence analysis of exon 7 from an R329X heterozygote, compared with a normal control. (C) PCR induced TaqI restriction site assay of LDLR gene codon 329 analysed by microplate array diagonal gel electrophoresis (MADGE). A series of probands and family members (affected and unaffected) is included. PCR product from unaffected subjects is 52 base pairs and digests to completion to yield a single band containing two overlying 26 base pair fragments. In R329X heterozygotes, the PCR product from the X allele contains the sequence TTGA instead of TCGA and therefore does not digest with TaqI; in consequence undigested PCR product as well as some digested PCR product is observed in R329X heterozygotes. In the MADGE system, this is observed as a two band instead of one band pattern.
Figure 2  Microsatellite D19S394 flanking the LDLR gene. (A) Genotype analyses of variable number tetranucleotide repeat D19S394 in a selection of affected members from kindreds carrying the LDLR R329X mutation. Each affected member bears an allele for marker D19S394 of 259nt, and there is a wide range of different sizes representing heterozygosity greater than 0.9 at this locus. Sizes estimated from interpolation from size markers are assigned to the nearest tetranucleotide length and are usually ±1 nt from the size, for example, all the 259 alleles shown here are estimated to be 259-260nt. Relative fluorescence intensity is marked on the scale on the right hand side. Small additional peaks frequently occur with this PCR, typically including also one small “stutter” peak 4nt shorter than the real allele size, but the true allele sizes are clearly readable. (B) Example of one small branch of one of the R329X kindreds studied, showing cosegregation of FH phenotype, R329X mutation, and allele size 259nt of D19S394. TC = total cholesterol (mmol/l), years of age, for example, 75 y. D19S394 genotypes are given as allele sizes in nucleotides, for example, 259, 267. (C) Distribution of allele sizes for D19S394 for a control group from the south of England.

disequilibrium between R329X and D19S394 allele size 259nt, all R329X probands displaying this allele size, although it only occurs in 16.1% of control chromosomes. Additionally, it should be noted that D19S394 is a tetranucleotide repeat which amplifies well in PCR, is easy to size accurately and precisely owing to the four base spacing of alleles and absence of stutter bands (fig 2), and has consistently shown heterozygosity above 0.9 in all populations studied (fig 2) (L Haddad, I Day, in preparation).

Discussion
The identification of LDLR mutation R329X enables direct diagnostic testing for approximately 10% of FH kindreds in the south of England; prevalence elsewhere is unknown. However, it has previously been reported in a single Norwegian family from Fossum, although the affected grandparent from this isolated family originated from an area north of London, UK (T Leren, personal communication). An initial survey of one London clinic showed 2/50 families positive for R329X; one family is from Wales (J Betteridge, personal communication). R329X has also been observed in Holland, although the earlier ancestry is not yet known. It should be noted that the apparent FH ancestor of one family examined in this study originated in York. R329X occurs at a CpG site and might represent a recurrent mutation at a CpG “hotspot.” However, there is strong evidence (see below) that R329X is identical by descent in all of the patients studied here. These data, taken together, suggest that R329X from a remote common ancestor may be relatively widespread in the UK and possibly in Europe, but probably with local expansion in numbers in the south of England. As discussed previously, R329X is a null mutation; the characteristics of the families studied here are consistent in that all display clearcut FH.

Several approaches to construction of a simple direct genetic assay for R329X were examined. Sequencing is definitive but laborious. SSCP is somewhat simpler, but still introduces a complex stage of post-PCR analysis. Additionally, the SSCP band pattern of R329X cannot be considered definitive, although from a range of different mutations identified in exon 7 of the LDLR gene we have not observed any other mutation with an identical
Common LDLR gene mutation

SSCP profile. The R329X mutation creates a BsrI site (5'-CCAGT), opening the possibility of a restriction site assay completely specific to R329X. However, BsrI, a thermophilic restriction enzyme, is well known to have stringent requirements for activity, and it proved necessary to purify PCR product (and concentrate it) to enable sufficient digestion and product detectable using ethidium bromide. Although fully specific, this procedure proved no more convenient than SSCP and was therefore used solely for verification of probands. Use of a PCR induced restriction site, such that an artificially created TaqI site is present in the PCR product from normal alleles, but is eliminated by the R329X mutation, proved to be simple and convenient. Although any base change in the CpG of codon 329 would eliminate the TaqI site, the assay is simple and sensitive (that is, not prone to false negatives, normal PCR product is digested completely), and can be reinforced where needed by BsrI assay, SSCP, or sequencing. We have since extended this approach\(^1\) to a PCR system applicable to any CpG site, in which both PCR primers are situated directly adjacent to the CpG dinucleotide of interest and both primers have a 3' T base (forcing if necessary), so that in the normal state a TaqI site is induced. If the CpG site contains a mutation, the TaqI site is lost. This system (CpG-PCR) creates a final common pathway for the analysis for any CpG mutation and further details of a range of CpG-PCR induced restriction site assays examined for R329X, as well as several other mutations, have been presented.\(^1\)

The observation that R329X is in complete linkage disequilibrium with allele 259nt of D19S394 in the probands from the south of England has several important implications. Firstly, within a collection of probands from a particular region, it may be possible to use D19S394 to identify subset(s) with a common mutation. Whether formal statistical significance is attained will depend on the number of FH probands in the study, the frequency in a control group of the D19S394 allele size marking the common mutation, and the frequency of the common mutation in the FH probands. For 100 probands, a mutation at 10% prevalence and marked by a D19S394 allele size with 10% population prevalence, 30 instead of 20 out of 200 alleles will have that allele size. For a mutation accounting for 20% of FH, the figures would be 40 instead of 20 out of 200 alleles. The respective chi-square values are 2.28 (p approximately 0.1) or 7.84 (p<0.01). It would be expected, therefore, that in smaller or less complex populations than the UK, and in regions with stable populations and where local expansion of a particular mutation may have taken place, this approach will prove useful. For example, in Norway where the population is approximately 4 million, intron 3 mutation 313+1 C>A ("Silverum") occurs in 25% of FH probands.\(^2\) Founder mutations have been reviewed previously in other populations such as Afrikaners, Finns, Christian Lebanese, and French-Canadians.\(^3\) Secondly, if a subset in linkage disequilibrium with an allele of D19S394 is identified, then this allele size will cosegregate with FH in the relatives of these probands. In the south of England, this second implication means that the apparent excess of D19S394 allele size 259nt marked a common FH mutation. Thirdly, since the population frequency of allele 259nt of D19S394 is only 16.1%, the likelihood of all R329X probands sharing this allele would be extremely small if R329X were a recurrent mutation. Rather, the conclusion is that all R329X probands are identical by descent, which is why the linkage disequilibrium approach can be used at all. D19S394 is located 250 kb 5' to the LDLR gene,\(^22\) which would be estimated to represent 0.25 cM recombination distance. Thus, this is expected to be a 1/400 chance of recombination between loci D19S394 and LDLR in a single generation. It is therefore consistent that we have not been able to link the R329X families from the south of England, since their common ancestor could date back many generations. In 1000 years (estimated 50 generations), 88.2% of the initial complete linkage disequilibrium would be expected still to exist. At the level of five generations, six large kindreds, two representing merges of smaller kindreds, have been identified in Southampton and south west Hampshire. It would be surprising if their common ancestor were 5000 to 6000 years ago, because then 50% of R329X chromosomes would no longer be expected to show allele 259nt of marker D19S394. However, a common ancestor up to 1000 years ago would be plausible. Population movements on account of migrations, wars, etc, whatever the original locality of R329X, could also be consistent on this timescale with wider spread not only in the south of England but also to the geographical extremes so far observed, namely Holland and Wales.

In conclusion, the LDLR gene mutation R329X is a significant contributor to the burden of FH in the south of England. Of wider applicability in FH, the use of polymorphic microsatellite D19S394 has been illustrated for linkage disequilibrium studies and family studies. The use of flanking microsatellites in place of intragenic RFLP based haplotyping warrants further development for analysis by linkage disequilibrium and cosegregation of FH mutations both in populations and families.

---