

Sequencing of the coding exons of the LRP1 and LDLR genes on individual DNA samples reveals novel mutations in both genes

F. Van Leuven^{a,*}, E. Thiry^a, M. Lambrechts^a, L. Stas^a, T. Boon^a, K. Bruynseels^a,
E. Muls^b, O. Descamps^c

^a *Experimental Genetics Group (EGG), Center for Human Genetics (CME), Flemish Institute for Biotechnology, K.U.Leuven-Campus Gasthuisberg O&N 06, B-3000 Leuven, Belgium*

^b *Department of Endocrinology, UZ-Gasthuisberg, K.U.Leuven, Leuven, Belgium*

^c *Department of Internal Medicine, Hôpital de Jolimont, Haine Saint-Paul, Belgium*

Received 3 March 2000; received in revised form 26 July 2000; accepted 31 August 2000

Abstract

Five coding polymorphisms in the LRP1 gene, i.e. A217V, A775P, D2080N, D2632E and G4379S were discovered by sequencing its 89 exons in three test-groups of 22 healthy individuals, 29 Alzheimer patients and 18 individuals with different clinical and molecularly uncharacterized lipid metabolism problems. No genetic defect was evident in the LRP1 gene of any of the Alzheimer's disease (AD) patients, further excluding LRP1 as a major genetic problem in AD. Lipoprotein receptor related protein (LRP) A217V (exon 6) was clearly present in all groups as a polymorphism, while D2632E was observed only once in a healthy volunteer. On the other hand, LRP1 alleles A775P, D2080N, and G4379 were encountered only in patients with FH or with undefined problems of lipid metabolism. This finding forced one to also analyze the LDL receptor (LDLR) gene, for which a method was devised to sequence the entire region comprising LDLR exons 2–18. The resulting sequence contig of 33 567 nucleotides yielded finally an exact physical map that corrects published and listed LDLR gene maps in many positions. In addition, next to known mutations in LDLR that cause FH, four novel LDLR defects were defined, i.e. del e7–10, exon 9 mutation N407T, a 20 bp insertion in exon 4, and a double mutation C292W/K290R in exon 6. No evidence for pathology connected to the LRP1 'mutations' was obtained by subsequent screening for the five LRP1 variants in larger groups of 110 FH patients and 118 patients with molecularly undefined, clinical problems of cholesterol and/or lipid metabolism. In three individuals with a mutant LDLR gene a variant LRP1 allele was also present, but without direct, obvious clinical compound effects, indicating that the variant LRP1 alleles must, for the present, be considered polymorphisms. © 2001 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: LRP1 gene; LDLR gene; Lipoprotein receptor related protein; LDL receptor; Mutation analysis

1. Introduction

Lipoprotein receptor related protein (LRP) contains many sequence motifs that resemble those of the classic LDL receptor (LDLR) [1]. LRP is not only a receptor for chylomicron remnants and lipoproteins carrying apolipoprotein E (ApoE) [2] but is the most wide-spectrum endocytosis receptor known. At least 30 unrelated ligands are known from areas of lipid metabolism,

proteinase inhibition, haemostasis, as well as some bacteria and viruses. This array is structurally reflected in the large size of the LRP precursor of about 600 kDa, with three clusters of cysteine-rich domains, each structurally equivalent to the classic LDLR (for reviews see [3–7]).

Following the epidemiological implication of the ApoE4 allele in Alzheimer's disease (AD), which is undisputed, but also unexplained, LRP has been implicated in the pathogenesis of AD [8–10]. The normal physiological role of LRP in CNS is not known and, while normally expressed by neurons, its ectodomain is present in the amyloid plaques in AD brain [8,9,11]. The genetic association of AD to the LRP1 locus on

* Corresponding author. Tel.: +32-16-345888; fax: +32-16-345871.

E-mail address: fredvl@med.kuleuven.ac.be (F. Van Leuven).

chromosome 12q13 [12–14] has been disputed [15–18]. The recent association of AD to the A2M gene, also located on chromosome 12 [19], attracted attention because LRP was originally identified as the A2M-receptor [20], while A2M itself is present in amyloid plaques in AD brain [21]. It was not possible, however, to obtain evidence for the proposed link, neither by genetic nor by biochemical analysis [22].

Understanding the genetic or functional effect of the ApoE4 allele or protein in the pathogenesis of AD, eventually including LRP, has instigated an intense exploration of these molecules for functional aspects in CNS lipoprotein metabolism. Experimental mouse models lacking LRP, LDLR, ApoE, A2M, the LRP-associated protein, the amyloid precursor protein, have all failed to be directly informative, either because of phenotypic silence with respect to CNS [23–30] or especially for LRP, because of embryonal lethality [31]. The induced mutant animal models have not yet provided insight into the physiology of LRP in adult and aging brain.

Human genetics, with its experiments of nature, offers the alternative to induced mutants. This has been very informative for understanding the LDLR gene and protein, and its function in patients with familial hypercholesterolemia (FH). This was helped considerably by the relatively direct receptor-ligand relation and by the many and diverse mutations in the LDLR gene, all resulting in the same defined clinical picture. That situation is completely opposite to what can be expected for mutations in the LRP gene: mutations could affect any of the many ligands and systems summed up above. In addition, the scarcity of genetic data on the human LRP1 gene contrasts sharply with the wealth of information on the LDLR gene. To help break this vicious circle, a physical map of the human LRP1 gene [32] and a method to sequence its 89 exons on individual DNA samples were composed [33].

Now the 89 exons and boundaries of the LRP1 gene have been analyzed in a group of 29 AD patients, without finding a consistent genetic defect, further extending the growing notion that LRP1 is not a prime genetic cause for AD [15]. Analyzed a heterogeneous group of 18 individuals with different clinical problems of lipid metabolism were analyzed, for which no molecular diagnosis was available or considered at the time of the analysis. This yielded novel mutations in the LRP1 gene in several individuals with high plasma cholesterol and lipid levels, which evidently necessitated also analysis of their LDLR gene, being the first candidate genetic defect [34]. Therefore a similar strategy was devised that was used for the LRP1 gene using long-range PCR and sequencing of exons 2–18 on individual DNA samples. The method identified carriers of known LDLR mutations and deletions, providing molecular diagnosis of FH. To the authors' surprise, this rela-

tively small group of 18 patients also contained four novel defective LDLR alleles, not listed nor reported before. The study was then further expanded to screening a large group of patients directly for the five LRP1 variant alleles. This identified only three individuals that carried a combination of a variant LRP1 gene and a mutant LDLR gene. The apparent lack of a compound disease effect in these individuals, although requiring further investigation, was interpreted to mean that these LRP1 variants are to be considered polymorphisms.

In addition to development of methods and characterization of novel mutations in both LRP1 and LDLR genes, a sequence contig of 33 567 nucleotides was composed that spans the entire genomic region from exon 2 to 18 of the human LDLR gene. This finally yielded the correct size of all but one intron, many at variance with published or listed data, while highlighting the incredible number of sequence repeats, mostly of the Alu-family, that are embedded in all the intervening sequences of the human LDLR gene.

2. Materials and methods

2.1. Individuals and patients

Genomic DNA was obtained from informed healthy individuals and from patients with informed consent or from responsible relatives, in case of children or elderly. The procedures were performed conforming with all existing ethical regulations.

The three initial test-groups consisted of 22 healthy individuals, 29 Alzheimer patients and 18 individuals with different clinical and molecularly uncharacterized lipid metabolism problems. The 22 healthy volunteers were in good health, without medical problems themselves or in their families, regarding lipid metabolism or cardiovascular function.

The AD patients were diagnosed by accepted standard criteria as described [8,13,18,22].

The initial group of 18 patients, for whom both the LDLR and LRP1 genes were fully analyzed by sequencing, consisted of a random selection from the university hospital out-patient lipid clinic, presenting with molecularly uncharacterized problems of high cholesterol and/or triglyceride levels in plasma. For screening, this group was extended and denoted as 'undefined' since they were a very heterogeneous group of patients that suffered diverse clinical problems associated with high plasma cholesterol and/or lipid levels, i.e. possible FH, hyperlipidemia and hypertriglyceridemia of different types, but without any molecular analysis. The 110 FH patients were authenticated (Table 1) as all had total plasma cholesterol levels higher than 300 mg/dl and a proven familial history of xanthomata

and/or early myocardial infarction, while most (85%) carried LDLR mutations.

Only ranges of values are given (Table 1) because of the heterogeneity of all groups and of the clinical symptoms and in biochemical data, which were further confounded by different types of medication and diet, given or installed before these patients presented at the university hospital. This situation reflects evidently the obvious problem, clinical and diagnostic, biochemical and molecular, which confronts the academic clinician facing this diverse group of patients. The present study was aimed as a first approach to analyze this diversity by searching for markers or mutations in the LRP1 gene, for the biological reasons explained in Section 1.

2.2. Analysis of the LRP1 gene

The methods to amplify the region of about 85 kb of the LRP1 gene from individual genomic DNA, using 14 long-range PCR reactions and subsequent sequencing of the 89 exons using 114 sequencing primers, was detailed previously [33]. Per reaction, about 20 ng purified genomic DNA was amplified in 50 µl reaction cocktails using a commercially available system (Expand System, Roche & Boehringer-Mannheim, Germany). All amplification programs were of the 'touch-down' type, i.e. after denaturation by heating at 94°C, four trains of three cycles each were performed with decreasing annealing temperatures between trains, followed by 30 cycles at the lowest temperature. Each cycle consisted of three steps: 30 s denaturation at 94°C, 1 min annealing at the indicated variable temperature, and 10 min extension at 68°C, with a final extension at 68°C for 15 min and subsequent cooling to 4°C (Geneamp 9700, PE-ABI Biosystems, Foster City, CA).

Amplicons were purified (Qiaquick procedure, Qiagen, Germany) and used as templates for Big Dye terminator cycle sequencing (FS sequencing; PE-ABI Biosystems, Foster City, CA). Between 50 and 200 ng template DNA was used per reaction, 4 µl of Big Dye terminator mix and 1 µM of sequence primer, in a final volume of 20 µl. Cycle sequencing consisted of denaturation for 3 min at 96°C, followed by 35 cycles of 10 s

each at 96°C, annealing for 5 s at a temperature ranging from 55 to 64°C depending on the melting temperature of the primer, and extension at 60°C for 4 min. The sequence reaction products were collected by centrifugation after ethanol-precipitation and dissolved in 25 µl template suppression reagent (PE-ABI Biosystems, Foster City, CA), denatured for 2 min at 95°C and analyzed by capillary electrophoresis using proprietary polymers and running buffer formulations, exactly as recommended by the manufacturer (ABI-Prism 310, PE-ABI Biosystems, Foster City, CA).

Screening for the novel LRP1 mutations in a larger group of individuals and patients was performed by simplified assays based on classic PCR amplification of individual genomic DNA to yield small amplicons for restriction analysis as specified (Table 2, Fig. 1).

2.3. Analysis of the LDLR gene

Similar general procedures were used to amplify and sequence the coding exons of the LDLR gene. A previously published method [35] was considerably simplified and optimized, resulting in three amplicons that cover the human LDLR gene from exon 2 to 18, used for sequencing. Three primer sets were used for long-range amplification and 17 sequencing primers for the coding exons 2–18 of the LDLR gene (Tables 3 and 4, Fig. 2). Separate amplification and sequence primer sets were developed for the promoter region including exon 1 of the LDLR gene. Methods for long-range amplification of the LDLR gene from individual DNA samples, isolation of amplicons and cycle sequencing reactions were similar to those developed and described for the LRP1 gene [33]. Selected amplicons or restriction fragments were subcloned and sequenced with the Genome Priming system, as instructed by the manufacturer (GPS system, New England Biolabs, Beverly, MA). Standard methods and procedures for molecular analysis were as described [32,33,36]. Sequence data were compared, corrected and assembled with dedicated software packages (Sequence Navigator and Factura, PE-ABI Biosystems, Foster City, CA; Lasergene, DNA-star, Madison, WI).

Table 1
Different groups of individuals and patients analyzed

Individuals	<i>n</i>	Age (years)	Sex (M/F)	Total plasma cholesterol	Total plasma triglycerides
Healthy	22	24–52	10/12	ND	ND
AD	29	63–84	12/17	ND	ND
FH	110	6–62	58/62	> 300	ND
'Undefined'	118	18–75	58/62	210–370	134–322

Table 2
Details of the screening method for the five mutations in the LRP1 gene based on PCR and restriction

Mutation	Exon	DNA Sequence	PCR primers code	Sequence	<i>n</i>	Amplicon (bp)	Enzyme	Fragments
A217V	6	GGGCC GGGTCC	L593 (999S) NE071 (1285A)	GATACCTCCTGCAGCCGAT GCAGACTGAGGGAGATGTTG	20 20	431	ApaI	244+187 431
A775P	14	GGCG GGCC	L417 (2721S) L416 (2881A)	GCAACTACCTCTTCTGGACT CATTGTTACCCCGGCATTTG	20 20	519	HaeIII	104 67+37
D2080N	39	TCGA	L534 (6484S)	TCTAGACAAGCCCCGGGCCA	20	979	TaqI	712+267
		TCAA	L536 (6861A)	GATGCCGGTTCGCAGGGGCA	20			979
D2632E	48	GACGCC	L879 (8275S)	AGCCTGTGGTGTGGGCGAGT	20	518	BsaHI	87+31
		GAAGCC	L338 (8431A)	GCTCGCAGGGCTGGAAGAGC	20			118
G4379S	85	CGGCCACTG	L139 (13484S)	GAGGTGAACAAGTGCAGCCG CTGT	24 20	516	AlwNI	323
		CAGCCACTG	L659 (13685A)	AAGACGTGCTCCTCACACCG	20			258+65

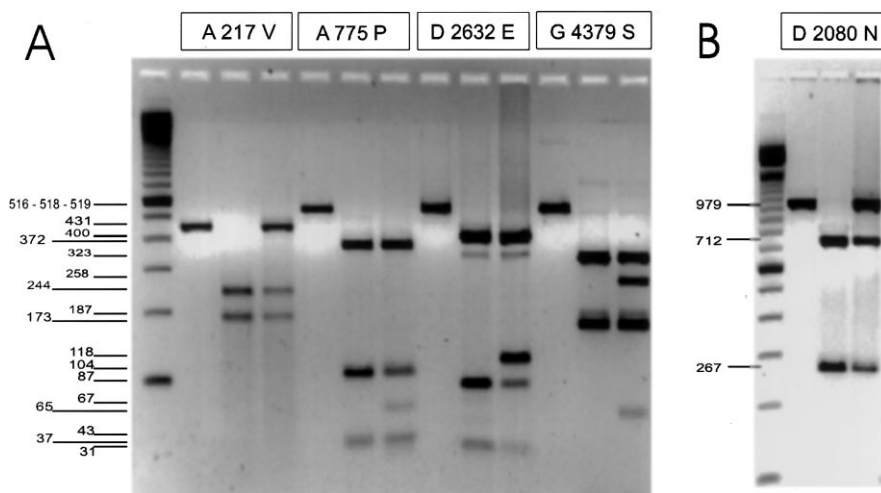


Fig. 1. Analysis for the five point mutations in the LRP1 gene by polymerase chain reaction (PCR) amplification and restriction. Gel-electrophoresis of small amplicons obtained by PCR with specific primers and restricted with specific enzymes (see Table 1 for details). For each mutation three samples are shown: (i) the specified LRP1 amplicon; and (ii) restricted from wild-type; and (iii) restricted from mutant LRP1. Left lanes contain DNA markers (100 bp). The size (in base pairs) of each fragment is indicated on the left of each gel. Electrophoresis was in 4% agarose (panel A) and in 2% agarose (panel B) (see also Section 2).

Table 3
Primers for long-range amplification of the LDLR gene

Amplicon	Size (bp)	Primer		Primer sequence	<i>n</i>	PCR program
		Code	Position			
ex 2–4	5263	NE290	i1S	5'-CCT TTC TCC TTT TCC TCT CTC TCA G	25	1
		NE295	531A	5'-CGC TGC GGC CAC TCA TCC	18	
ex 4–11	10 763	NE294	446S	5'-GTC CCG CCA GCT TCC AGT GC	20	2
		NE291	1588A	5'-TTC CCC AGT CAG TCC AGT ACA TGA A	25	
ex 10–18	18 649	NE292	1562S	5'-CCA TCG TGG TGG ATC CTG TTC ATG G	25	3
		NE293	3679A	5'-CCT CAC CGT GCA TGT TTT AAA CAC A	25	

Table 4
Primers for sequencing exons 2–18 of the LDLR gene

Code	Position	Sequence	<i>n</i>
NE589	i2A	5'-CCA ATA ATT CAT TCT CTC CC	20
NE330	i3A	5'-CCC CAG GAC TCA GAT AGG	18
NE294	446S	5'-GTC CCG CCA GCT TCC AGT GC	20
NE595	i3S	5'-CAC CTA TTA GCG CAC CAG TG	20
NE331	i4S	5'-AGG CCC TGC TTG TTT TTC T	19
NE303	i5S	5'-CAC ACC TGA CCT TCC TCC TT	20
NE332	i6S	5'-GGC GGC GAA GGG ATG GGT AG	20
NE338	i7S	5'-GAA GAG CCT CCC CAC CAA	18
NE305	i8S	5'-TCC CCT CTG ACC CCC TGA C	19
NE333	i9S	5'-TCC TGG TGC GAT GCC CTT CTC T	22
NE334	i10S	5'-CCG CCC TCC AGC CTC ACA G	19
NE306	i12A	5'-ACC TCC TCC TAG TCA CAA CC	20
NE307	i14A	5'-GGA GGA CAC AGG ACG CAG AA	20
NE335	i14S	5'-GGC GCA CAC CTA TGA GAA G	19
NE336	i16A	5'-CCA GGC ACG AGG TCA CAT AG	20
NE337	i16S	5'-AGC TGG GTC TCT GGT CTC G	19
NE594	2614A	5'-GCG AGG TCT CAG GAA GGG TT	20

3. Results

3.1. Mutation analysis of the LRP1 gene by sequencing

The LRP1 gene was analyzed by a combination of long-range PCR and sequencing [32,33] in three test-groups of 22 apparently healthy individuals, 29 Alzheimer patients and 18 individuals presenting with different clinical problems of cholesterol and lipid metabolism of unknown etiology. This analysis identified several hundreds of DNA polymorphisms scattered all over the exons and introns of the LRP1 gene (results not shown) in addition to those previously reported [33].

More importantly, besides the A217V polymorphism in exon 6 of the LRP1 gene [33], four novel exonic point mutations were identified predicting variant amino acids in the LRP precursor protein (Table 5). The mutations are denoted by the original and mutant amino acid, separated by its position in the LRP precursor protein translated from the human LRP cDNA designating methionine corresponding to of the ATG start codon as amino acid 1 [1](human LRP cDNA, Genbank accession number X13916). Exons were numbered and defined on the physical map of the LRP1 gene, published with accession numbers of partial sequences deposited in Genbank [32,33].

3.2. Screening for the five LRP1 mutations

Subsequent screening of larger groups of individuals and patients for the five LRP1 mutations was performed by simplified assays based on PCR amplification of individual genomic DNA to yield small amplicons for restriction analysis (Table 2, Fig. 1). A

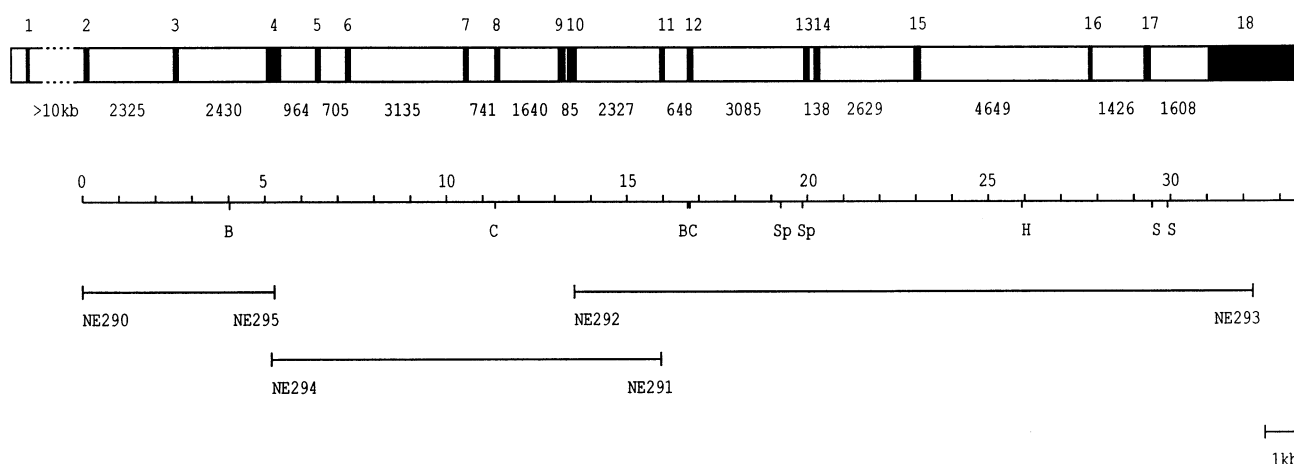


Fig. 2. Physical map of the LDL receptor (LDLR) gene based on the new sequence contig. The exons are numbered and represented by the filled blocks, the size of the intervening sequences is indicated in base pairs under each intron. The sequence contig of 33 567 nucleotides is represented by the scale divided in kilo-base pairs, indicated by the numbers above it. The three overlapping long-range amplicons are schematically represented with the primers indicated (see also Table 3). Rare or relevant restriction sites are indicated: B, Bgl II; C, Cla I; Sp, Spe I; S, Sst I; H, HindIII.

Table 5
Occurrence of the the LRP1 mutations in different groups of individuals

Individuals ^a	N	Ex 6 A217V	Ex 14 A775P	Ex 39 D2080N	Ex 48 D2632E	Ex 85 G4379S
Healthy	22	2	0	0	1	0
AD	29	1	0	0	0	0
FH	110	4	1	3	0	1
'Undefined'	118	4	4	2	0	1
Total	279	11	5	5	1	2

^a See Table 1 and text for details.

total of 558 alleles were analyzed and the five LRP1 mutations, i.e. A217V(exon 6), A775P(exon 14), D2080N(exon 39), D2632E(exon 48) and G4379S(exon 85) were encountered in respectively 11, five, five, one, and two instances (Table 5). The 24 mutant LRP1 alleles represent 4.3% of alleles analyzed, which is higher than expected, given the highly conserved nature of the LRP1 cDNA and gene [1,32,33].

In the cohort of 29 AD patients, originating from different centers, only a single copy of the A217V mutation was detected, while observed in two healthy persons (Table 5). Eventual physiological repercussions are unlikely, corroborating a previous sequencing effort of the coding exons of the LRP1 gene in a group of 12 selected AD patients [18]. The combined results strengthen the conclusion that the LRP1 gene is not likely to be a primary genetic problem in AD [18].

A more obvious indication for analyzing the LRP1 gene for mutations was in patients with problems in cholesterol and lipid metabolism. Among a total of 228 such patients analyzed for the five LRP1 mutations, eight alleles were detected with mutation A217V, five alleles with mutation A775P, five alleles with mutation D2080N, no alleles with mutation D2632E and two alleles with mutation G4375S (Table 5).

Mutation A217V was encountered most frequently, with 11 occurrences in 279 individuals analyzed, but was evidently not related to disease, at least not when present in the heterozygous state. One expects to eventually encounter homozygous carriers of this mutation, which would allow one to decide for or against a physiological repercussion of the LRP1 A217V mutation. Mutation D2632E was detected only once and in a healthy person. The three other LRP1 mutations were all observed in patients suffering from different lipid metabolism problems of unknown etiology. Further analysis was initiated by analyzing their LDLR gene.

3.3. Sequencing of the LDLR gene on individual DNA samples

The original, heterogeneous group of 18 individuals presenting with diverse clinical problems of increased

cholesterol and/or triglyceride levels, were never analyzed molecularly before nor was such analysis planned. To allow a precise molecular diagnosis and to allow identification of eventual compound effects of the LRP1 mutations, a sequencing strategy was devised for exons 2–18 of the LDLR gene. Efficient long-range amplification of individual DNA samples was obtained with three primer sets producing amplicons covering the region from exon 2 to 18 (Fig. 2, Table 3). The size and the complexity of intron 1 prevented its efficient amplification, and analysis was not further pursued. A set of 17 synthetic oligonucleotides was composed to sequence exons 2–18 and their boundaries (Table 4), making nearly complete analysis of the LDLR gene amenable and efficient, and evidently much simpler than analysis of the LRP1 gene. Analysis of the promoter region including exon 1 of the LDLR gene was also implemented, although not further detailed here, by including an extra PCR primer set yielding an amplicon of about 0.8 kb that could be sequenced using two extra primers.

3.4. Identification of known and novel LDLR mutations

In five of the 18 'undefined' patients in which the entire LDLR gene was sequenced, mutations known or suspected to cause FH [37–41] (LDLR database at www.ucl.ac.uk/fh) were observed, i.e. E10X, C122X, A370T, and a 3 kb deletion in two unrelated persons, comprising exons 7 and 8 as further detailed below. This molecularly confirmed the clinical diagnosis of FH.

Rather to the authors' surprise in this small group of 18 patients, four individuals were molecularly established as carriers novel defects in the LDLR gene, not characterized or reported before.

First, a mutation N407T in exon 9 was predicted from a change of A to C at position 1283 of the LDLR cDNA.

Second, an insertion of 20 bases at position 681, i.e. after codon 206 in exon 4, which is different from the 21 bp insertion known as FH Tulsa-1 [37,42]. Although

both insertions occur at the same site, the novel 20 bp insertion is predicted to disturb the reading-frame and yield a truncated LDLR receptor, while FH Tulsa-1 caused duplication of codons 200–206 in the LDLR protein [42].

Third, a patient originally thought to be a compound heterozygote for mutations C292W and K290R, was actually carrier of a double mutant allele, as proven by analyzing its relatives (results not shown). The C292W and K290R mutations have been listed in the databases, but have apparently never been observed on the same allele.

Finally, an individual was found to carry a large deletion of about 3.5 kb, different from the 3 kb deletion, as detailed below.

3.5. Characterization of LDLR del e7–8

The two LDLR alleles that contained large deletions of 3 and 3.5 kb were readily distinguished by agarose electrophoresis of the long-range PCR amplicons (Fig. 3A). Subsequent sequencing was negative for exons 7 and 8 in both amplicons and, additionally for exons 9 and 10 in the 3.5 kb deletion. Molecularly, the 3 kb deletion was nearly identical to del e7–8, previously reported in Dutch and Belgian FH patients ([40], and references therein). Amplification and sequencing of the fusion-site demonstrated, however, differences in the T-rich repeat that has been proposed to cause the deletion [40]. Now stretches of 15 or 16 T-nucleotides were identified in unrelated persons, as opposed to the 12 T-nucleotide stretch previously reported [40].

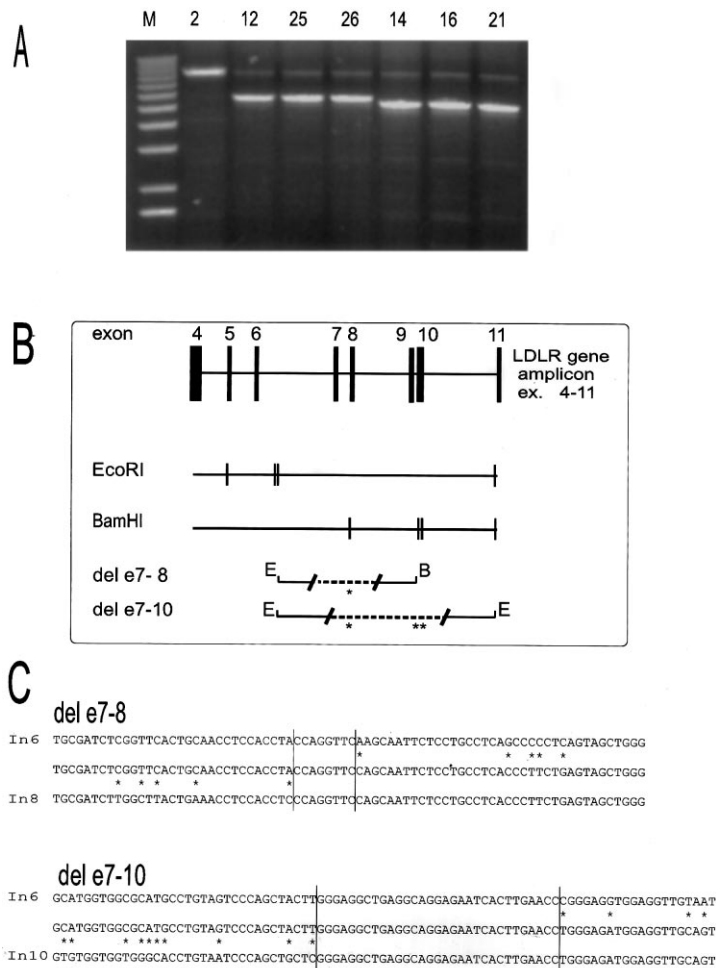


Fig. 3. Analysis of two large deletions in the LDL receptor (LDLR) gene. Panel A: gel electrophoresis (0.7% agarose, ethidium bromide staining) of long-range polymerase chain reaction (PCR) amplicons of the LDLR gene, wild-type (lane marked 2), del e7–8 (lanes marked 12, 25, 26) and del e7–10 (lanes marked 14,16, 21). The left lane contains 1 kb markers. Note in each lane also the presence of the wild-type amplicon, which is less efficiently amplified when a shorter deleted allele is present. Panel B: schematic representation of the amplicon corresponding to exons 4–11 of the LDLR gene. Exons are marked by the filled blocks and numbered. Relevant EcoRI and BamHI restriction sites are indicated and can be on the LDLR sequence contig. The large deletions are represented on the restriction fragments (see text for details) by the broken lines. The asterisks denote the missing BamHI restriction sites in the deletions. Panel C: Alignment of the sequences of the fusion-points of del e7–8 and del e7–10 with the relevant intronic sequences of introns 6, 8 and 10 as indicated. Asterisks denote differences in the aligned sequences. The vertical lines indicate the area of sequence identity, between which the fusion must be located. Both areas are located in Alu repeat sequences (see text for details).

Table 6
Size and contig coordinates of exons of the LDLR gene and corrected intron size

Exon #	Exon size (bp)	Contig coordinates	Intron size (bp)
2	123	26–148	2325
3	123	2474–2596	2430
4	381	5029–5409	964
5	123	6374–6496	705
6	123	7201–7323	3135
7	120	10 459–10 578	741
8	126	11 320–11 445	1640
9	172	13 086–13 257	85
10	228	13 343–13 570	2327
11	119	15 898–16 016	648
12	140	16 663–16 802	3085
13	142	19 888–20 029	138
14	153	20 166–20 318	2629
15	171	22 948–23 118	4649
16	78	27 768–27 845	1426
17	158	29 272–29 429	1608
18	2530	31 038–33 567	–

The exact location of the fusion point was now determined by subcloning and sequencing a 2.2 kb EcoRI–BamHI restriction fragment from the long-range amplicon of the deleted allele (Fig. 3B). This fragment contained exon 9 and part of exon 10, but the junction or fusion-point was located differently than proposed [40]. Only with complete sequencing of the LDLR gene as described below, was alignment possible of the sequences of del e7–8 with introns 6 and 8. This proved that the Alu type repeats on both sides of the actual fusion site were very similar and defined in fact a ‘fusion-region’ or overlap of eight nucleotides (Fig. 3C). The first base of this octet is located 1528 bases downstream of the EcoRI site in intron 6, while its last base is 474 nucleotides upstream of the intron 8–exon 9 boundary (Fig. 3). From the LDLR gene sequence contig (see below), the deletion was unequivocally identified to correspond to bases 9600–12611 in that contig, with the caveat of an identical 8 bp overlap or margin at each border.

3.6. Characterization of LDLR del e7–10

The larger deletion of about 3.5 kb affected exons 7–10 as demonstrated by the negative outcome of sequencing for these exons, and sustained by Southern blotting (results not shown). From the literature and dedicated databases, only one instance was retrieved in which exons 7–10 appeared similarly deleted, in a single Finnish family with deletion ‘Oulu-1’. This deletion was 7.5 kb based on Southern blotting [43,44], while no molecular analysis nor DNA was available for comparative studies (K. Kontula, personal communica-

tion). It appears, although not proven, to be different from the current del e7–10 defect.

Characterization of del e7–10 by amplification, restriction analysis and partial sequencing, positioned the rearrangement in intron 7, immediately upstream of exon 7, and downstream of exon 10 (Fig. 3). Fusing of the larger part of introns 6 and 10 resulted in flanks of the fusion-point that were entirely intronic and unknown. Attempts to subclone an EcoRI fragment of about 4.5 kb from the deletion amplicon failed repeatedly, likely due to instability in bacterial cultures, as previously also encountered when intron 44 of the LRP1 gene was subcloned, that also contained long stretches of highly repetitive sequences [32].

The del e7–10 fusion point was defined on the amplicon by primer walking and shown to be positioned at between 124 and 93 bases upstream of the intron 6–exon 7 boundary (Fig. 3). The fusion site in intron 10, defined after obtaining the long sequence contig of the LDLR gene (see below), comprised nucleotides 10 366–13 837. Alignment of sequences of del e7–10 with introns 6 and 10 then proved this deletion to involve very similar Alu-type repeats to those in del e7–8, but with a larger overlap or margin of 31 identical nucleotides at each side (Fig. 3C).

3.7. Sequence contig of the LDLR gene

The effort to sequence the entire genomic region from exon 2 to 18 of the LDLR gene was instigated by the accumulated sequences resulting from the analysis as outlined above. In addition, the PCR and sequence data identified many inconsistencies with published or listed data. Moreover, during this study one became aware of a major diagnostic problem that remains associated with the molecular diagnosis of FH. Finally, a fundamental interest in the vast number of mutations that accumulate in the LDLR gene, can but be based on and profit from exact sequence information. To that end, the three large amplicons spanning the LDLR gene from exons 2 to 18 (Fig. 2) were subcloned and sequenced, making use of a random integration and sequencing system (GPS system, see Section 2). The resulting sequence contig of 33 567 nucleotides spanned the LDLR gene from position –25 relative to the intron 1–exon 2 boundary, and stretching up to the poly-adenylation signal in exon 18 (Fig. 2; Table 6).

Sequencing proceeded rapidly and uneventfully, except for six refractory sites for which the sequencing system did not yield readable data. Bridging these gaps was achieved by primer walking with synthetic oligonucleotides designed in both of the flanking regions. The last gap with severe compression was eventually solved using the dGTP version of the sequencing system. The final sequence contig of 33 568 nucleotides (Genbank accession number AF217403) contained only two unre-

solved sites at positions 1296 and 28 689–28 691. This finally resulted in an exact physical map, with exact coordinates of the exons on this contig and with exactly sized introns, which corrects in many positions, all previously published and listed coordinates of the LDLR gene (Fig. 2, Table 5). This will undeniably help to more precisely determine the nature and the location of molecular defects, as demonstrated and outlined above in characterizing the two large deletions (Fig. 3).

Very conspicuously, the LDLR contig contained an abundance of repetitive sequences, mainly of the Alu-family, embedded in the intervening sequences, as illustrated by a self-alignment dot-plot (Fig. 4). The many direct and inverted repeats practically leave only the exons to crop up as being unique sequences in the human LDLR gene (Fig. 4).

4. Discussion

The previous and current data do not sustain the hypothesis that the LRP1 gene would be directly involved in the pathogenesis of dementia of the Alzheimer type. No consistent mutation was found in its coding region in autopsy-proven late onset AD-patients, which leaves only the promoter region of LRP1 or its neighboring genes on chromosome 12 as candidates for further investigation. Nevertheless, the available molecular and epidemiological data do not provide a major incentive to launch studies along these lines in the field of Alzheimer research.

The first area of importance of these receptors thus remains in the field of cholesterol and lipid metabolism, further warranting their nomenclature. The current

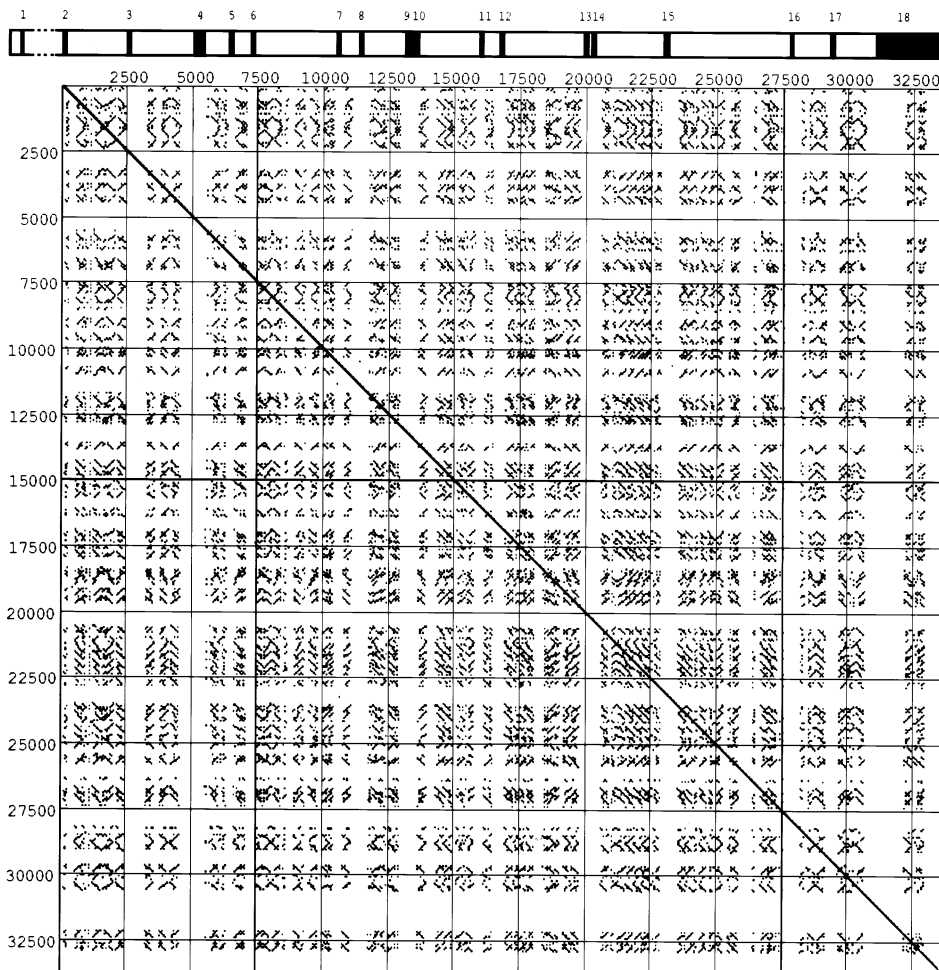


Fig. 4. Dot-plot of the sequence contig of the LDL receptor (LDLR) gene. To illustrate the repeat sequences, the 33 567 nucleotides of the LDLR sequence contig was compared to itself (Megalign, Lasergene, DNA-star). Each short diagonal line or fragment-line represents a direct or inverted repeat sequence. Note the 'sparing' of the exons, indicating these to be 'unique' sequences, while all intronic sequences contain repeats, mainly of the Alu family. Although most are partial, complete Alu repeats are obvious and represented by the ' ^ ' aspect of the line fragments in the figure.

study cannot but be considered an initial genetic step in the direction of elucidating the role of LRP1 in humans. In the absence of suitable model systems, i.e. experimental inactivation of the LRP1 gene in mice was embryonically lethal [31], the physiological role of LRP in adults remains unknown. The finding of individuals with cholesterol and lipid metabolism problems, who carry a combination of mutations in their LRP1 and LDLR genes, appeared initially promising. This was, however, not borne out by the screening of a larger group of patients and individuals, neither in a cohort of proven FH patients, nor in a group of molecularly uncharacterized patients. No specific or consistent prominent or even minor clinical or patho-physiological indication was evident in carriers of any of the five LRP1 mutations. In addition, no compound effect of the LRP1 mutations in conjunction with the LDLR mutations could be discerned. The variant LRP1 alleles must therefore be considered polymorphisms, until proven otherwise by further analysis, requiring detailed clinical, biochemical and genetic stratification.

Since the different LRP1 alleles, neither on their own nor in addition to the patho-genetic defects in the LDLR gene, caused or aggravated a given pathology or physiological problem, the original question remains unanswered. Additional work on the LRP1 gene is required to define its fundamental, and eventual diagnostic relevance. Combination with data from related genes, including the LDLR gene, might provide clues to the fundamental and clinical questions. Next to sequencing analysis of individual LRP1 genes, micro-array type technology is implemented to decipher the complex interplay and genetic variation, and to understand its fundamental significance in normal and pathological conditions.

In addition to the methodology to analyze the LRP1 and LDLR genes on individual DNA samples, novel mutations are reported in both genes. Moreover, a sequence contig was composed of 33 567 nucleotides that spans the genomic region from exon 2 to 18 of the human LDLR gene. This yielded finally the correct size of all introns, except intron 1, exposing many variations within previously published and listed data. The sequence contig further highlighted the incredible number of repeats, mostly of the Alu-family, present in the intervening sequences in the LDLR gene, which are evidently largely to blame for the accumulation of deletions and mutations in this 'morbid' gene. This is in clear contrast with the LRP1 gene, which maintains its status as a very conserved protein and gene [1,32,33,45,46], with a conserved and essential function [7,31], still waiting for further exploration in adult tissues.

Acknowledgements

We thank G. Matthys and the staff of molecular diagnostics of the Center for Human Genetics. We acknowledge the help of many scientists and doctors who contributed with advice and materials: A. Peeters (Stellenbosch, South Africa), L.F. Van Gaal (Antwerp, Belgium), L. Havekens, F. de Boer and M. Mohrschladt (Leiden, The Netherlands), J. Auwerx (Strasbourg, France), M.-C. Chartier-Harlin (Lille, France), J. Hardy (Jacksonville, Florida), W. Rebeck and B. Hyman (Boston, Massachusetts), P. St. George-Hyslop (Toronto, Canada), F. Heller (Hôpital de Jolimont, Belgium). This investigation was supported by the "Fonds voor Wetenschappelijk Onderzoek" (FWO-Vlaanderen), the Interuniversity-network for Fundamental Research (IUAP), the special Action Program for Biotechnology of the Flemish government (VLAB/IWT, COT-008), the Rooms Fund, and by K.U.Leuven R&D.O.D. is a fellow of the Belgian Heart League and the Belgian Lipid Club and was supported by the Hôpital de Jolimont.

References

- [1] Herz J, Hamann U, Rogné S, Myklebost O, Gausepohl H, Stanley KK. Surface location and high affinity for calcium of a 500-kDa liver membrane protein closely related to the LDL-receptor suggest a physiological role as lipoprotein receptor. *EMBO J* 1988;7:4119–27.
- [2] Beisiegel U, Weber W, Ihrke G, Herz J, Stanley KK. The LDL-receptor-related protein, LRP, is an apolipoprotein E-binding protein. *Nature* 1989;341:162–4.
- [3] Brown MS, Herz J, Kowal RC, Goldstein JL. The low density lipoprotein receptor-related protein: double agent or decoy? *Curr Opin Lipidol* 1991;2:65–72.
- [4] Krieger M, Herz J. Macrophage scavenger receptors and LDL receptor-related protein (LRP). *Annu Rev Biochem* 1992;63:601–37.
- [5] Moestrup S. The α_2 -macroglobulin receptor and epithelial glycoprotein-330: two giant receptors mediating endocytosis of multiple ligands. *Biochem Biophys Acta* 1994;1197:197–213.
- [6] Strickland D, Kounnas M, Agraves W. LDL receptor-related protein: a multiligand receptor for lipoprotein and proteinase catabolism. *FASEB J* 1995;9:890–8.
- [7] Willnow TE. The low-density lipoprotein receptor gene family: multiple roles in lipid metabolism. *J Mol Med* 1999;77:306–15.
- [8] Rebeck G, Reiter J, Strickland D, Hyman B. Apolipoprotein E in sporadic Alzheimer's disease: allelic variation and receptor interactions. *Neuron* 1993;11:575–80.
- [9] Tooyama I, Kawamata T, Akiyama H, Moestrup SK, Gliemann J, McGeer PL. Immunohistochemical study of alpha 2 macroglobulin receptor in Alzheimer and control postmortem human brain. *Mol Chem Neuropathol* 1993;18(1–2):153–60.
- [10] Weisgraber K, Mahley R. Human apolipoprotein E: the Alzheimer's disease connection. *FASEB J* 1996;10:1485–94.
- [11] Laferla FM, Troncoso JC, Strickland DK, Kawas CH, Jay G. Neuronal cell death in Alzheimer's disease correlates with apoE uptake and intracellular A stabilization. *J Clin Invest* 1997;100(2):310–20.

- [12] Kang D, Saitoh T, Chen X, et al. Genetic association of the low-density lipoprotein receptor-related protein gene (LRP), an apolipoprotein E receptor, with late-onset Alzheimer's disease. *Neurology* 1997;49:56–61.
- [13] Wavrant-De Vrieze F, Pérez-Tur J, Lambert J-C, et al. Association between the low density lipoprotein receptor-related protein (LRP) and Alzheimer's disease. *Neurosci Lett* 1997;227:68–70.
- [14] Lendon C, Talbot C, Craddock N, Han S, Wragg M, Morris J, Goate A. Genetic association studies between dementia of the Alzheimer's type and three receptors for apolipoprotein E in a Caucasian population. *Neurosci Lett* 1997;222:187–90.
- [15] Fallin D, Kundtz A, Town T, et al. No association between the low density lipoprotein receptor-related protein (LRP) gene and late-onset Alzheimer's disease in a community-based sample. *Neurosci Lett* 1997;233:145–7.
- [16] Pericak-Vance M, Bass M, Yamaoka L, et al. Complete genomic screen in late-onset familial Alzheimer disease. *J Am Med Assoc* 1997;278:1237–41.
- [17] Clathworthy A, Gomez-Isla T, Rebeck G, Wallace R, Hyman B. Lack of association of a polymorphism in the low-density lipoprotein receptor-related protein gene with Alzheimer disease. *Arch Neurol* 1997;54:1289–92.
- [18] Wavrant-DeVrieze F, Lambert JC, Stas L, et al. Association between coding variability in the LRP1 gene and the risk of late-onset Alzheimer's disease. *Hum Genet* 1999;104:432–4.
- [19] Blacker D, Wilcox M, Laird N, et al. Alpha-2 macroglobulin is genetically associated with Alzheimer's disease. *Nat Genet* 1998;19:357–60.
- [20] Van Leuven F, Cassiman JJ, Van Den Berghe H. Primary amines inhibit recycling of 2M-receptors in fibroblasts. *Cell* 1980;20:37–43.
- [21] Van Gool D, De Strooper B, Van Leuven F, Triau E, Dom R. Alpha-2-macroglobulin expression in neuritic-type plaques in patients with Alzheimer's disease. *Neurobiol Aging* 1993;14:233–7.
- [22] Rogava EA, Premkumar S, Grubber J, et al. An alpha-2-macroglobulin insertion-deletion polymorphism in Alzheimer disease. *Nat Genet* 1999;22(1):19–22.
- [23] Piedrahita JA, Zhang SH, Hagan JR, Oliver PM, Maeda N. Generation of mice carrying a mutant apolipoprotein E gene inactivated by gene targeting in embryonic stem cells. *Proc Natl Acad Sci USA* 1992;89:4471–5.
- [24] Plump AS, Smith JD, Hayek T, Aalto-Setälä K, Walsh A, Rubin EM, Breslow J. Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell* 1992;71:343–53.
- [25] Ishibashi S, Brown MS, Goldstein JL, Gerard RD, Hammer RE, Herz J. Hypercholesterolemia in low density lipoprotein receptor knockout mice and its reversal by adenovirus-mediated gene delivery. *J Clin Invest* 1993;92(2):883–93.
- [26] Van Ree JH, Van den Broek WJAA, Dahlmans VEH, et al. Diet-induced hypercholesterolemia and atherosclerosis in heterozygous apolipoprotein E-deficient mice. *Atherosclerosis* 1994;111:25–37.
- [27] Li ZW, Stark G, Götz J, Rüllicke T, Müller U, Weissmann C. Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc Natl Acad Sci USA* 1996;93:6158–62.
- [28] Hofker MH, van Vlijmen BJM, Havekes LM. Transgenic mouse models to study the role of ApoE in hyperlipidemia and atherosclerosis. *Atherosclerosis* 1998;137:1–11.
- [29] Umans L, Serneels L, Overbergh L, Stas L, Van Leuven F. 2-Macroglobulin- and murinoglobulin-1 deficient mouse. *Am J Pathol* 1999a;155(3):983–93.
- [30] Umans L, Serneels L, Lorent K, Dewachter I, Tesseur I, Moechars D, Van Leuven F. Lipoprotein receptor-related protein in brain and in cultured neurons of mice deficient in receptor-associated protein and transgenic for apolipoprotein E4 or amyloid precursor protein. *Neuroscience* 1999b;94(1):315–21.
- [31] Herz J, Clouthier DE, Hammer RE. LDL receptor-related protein internalizes and degrades uPA-ApA-1 complexes and is essential for embryo implantation. *Cell* 1992;71:411–21.
- [32] Van Leuven F, Stas L, Hilliker C, et al. Structure of the gene (LRP1) coding for the human Alpha-2-macroglobulin/lipoprotein receptor related protein. *Genomics* 1994;24:78–89.
- [33] Van Leuven F, Stas L, Thiry E, Nelissen B, Miyake Y. Strategy to sequence the 89 exons of the human LRP1 gene coding for the lipoprotein receptor related protein: identification of one expressed mutation among 48 polymorphisms. *Genomics* 1998;52:138–44.
- [34] Südhof TC, Goldstein JL, Brown MS, Russell DW. The LDL receptor gene: a mosaic of exons shared with different proteins. *Science* 1985;228:815–22.
- [35] Rødningsen OK, Leren TP. Application of long range polymerase chain reaction in the study of the LDL receptor gene. *Scand J Clin Lab Invest* 1996;56:93–6.
- [36] Maekawa B, Cole T, Seip R, Bylund D. Apolipoprotein E genotyping methods for the clinical laboratory. *J Clin Lab Anal* 1995;9:63–9.
- [37] Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat* 1992;1:445–66.
- [38] Soutar A. Update on low density lipoprotein receptor mutations. *Curr Opin Lipidol* 1998;9:141–7.
- [39] Varret M, Rabes JP, Thiart R, et al. LDL-R Database (second edition): new additions to the database and the software, and results of the first molecular analysis. *Nucleic Acids Res* 1998;26:248–52.
- [40] Peeters AV, Van Gaal LF, du Plessis L, Lombardi MP, Havekes LM, Kotze MJ. Mutational and genetic origin of LDL receptor gene mutations detected in both Belgian and Dutch familial hypercholesterolemics. *Hum Genet* 1997;100:266–70.
- [41] Descamps OS, Hondekijn JC, Van Acker P, Deslypere J, Heller FR. High prevalence of a novel mutation in the exon 4 of the low density lipoprotein receptor gene causing familial hypercholesterolemia in Belgium. *Clin Genet* 1997;51:301–8.
- [42] Webb J, Sun X-M, McCarthy S, Neuwirth C, Thompson G, Knight B, Soutar A. Characterization of mutations in the low density lipoprotein (LDL)-receptor gene in patients with homozygous familial hypercholesterolemia, and frequency of these mutations in HF patients in the United Kingdom. *J Lipid Res* 1996;37:368–81.
- [43] Aalto Setälä K, Koivisto UM, Miettinen TA, et al. Prevalence and geographical distribution of major LDL receptor rearrangements in Finland. *J Int Med* 1992;231:227–34.
- [44] Kontula K, Koivisto UM, Koivisto P, Turtola H. Molecular genetics of familial hypercholesterolaemia: common and rare mutations of the low density lipoprotein receptor gene. *Ann Med* 1992;24:363–7.
- [45] Van Leuven F, Stas L, Raymakers L, et al. Molecular cloning and sequencing of the murine Alpha-2-macroglobulin receptor cDNA. *Biochim Biophys Acta* 1993;1173:71–4.
- [46] Nimpf J, Stefan S, Bilous PT, Schneider WJ. The somatic cell-specific low density receptor-related of the chicken. *J Biol Chem* 1994;269:1–8.