 © 1998 Oxford University Press

LDLR Database (second edition): new additions to the database and the software, and results of the first molecular analysis

Mathilde Varret[1], Jean-Pierre Rabés[1,2], Rochelle Thiart[3], Maritha J. Kotze[3], Heike Baron[4], Ana Cenarro[5], Olivier Descamps[6], Margit Ebhardt[7], Jean-Claude Hondelijn[6], Gert M. Kostner[8], Yasuko Miyake[9], Miguel Pocovi[5], Hartmut Schmidt[10], Helena Schmidt[8], Herbert Schuster[4], Manfred Stuhrmann[7], Taku Yamamura[9], Claudine Junien[1,2], Christophe Béroud[1] and Catherine Boileau[1,2,*]

[1]INSERM U383, Hôpital Necker-Enfants Malades, Université René Descartes, Paris V, 149-161 rue de Sèvres, 75743 Paris Cedex 15, France, [2]Laboratoire Central de Biochimie et de Gntique Moléculaire, CHU Ambroise Paré, 9 avenue Charles de Gaulle, 92104 Boulogne Cedex, France, [3]MRC Cape Heart Group, Division of Human Genetics, Faculty of Medicine, University of Stellenbosch, PO Box 19063, Tygerberg 7505, South Africa, [4]Humboldt-Universität zu Berlin, Virchow-Klinikum, Franz-Volhard-Klinik am Max-Delbrück-Centrum für Molekulare Medizin, Wiltergstraße 50, 13122 Berlin, Germany, [5]Departamento de Bioquimica y Biologia Molecular y Cellular, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain, [6]Groupe d'Etude du mtabolisme tumoral, division lipides, Hôpital de Jolimont, 179 rue ferrer, 7100 Haine Saint-Paul, Belgium, [7]Institut für Humangenetik, Medizinische Hochschule, Hannover, Carl-Neuberg-Straße 1, 30625 Hannover, Germany, [8]Institute of Medical Biochemestry at the University of Graz, Harrachgasse 21, Graz A-8010, Austria, [9]Department of Ethiology and Pathophysiology, National Cardiovascular Center Research Institute, 5-7-1 Fujishirodai, Suita, Osaka 565, Japan and [10]Department of Gastroenterology and Hepatology, Medizinische Hochschule, Hannover, Carl-Neuberg-Straße 1, 30625 Hannover, Germany

Received October 1, 1997; Accepted October 3, 1997

ABSTRACT

Mutations in the LDL receptor gene (LDLR) cause familial hypercholesterolemia (FH), a common autosomal dominant disorder. The LDLR database is a computerized tool that has been developed to provide tools to analyse the numerous mutations that have been identified in the LDLR gene. The second version of the LDLR database contains 140 new entries and the software has been modified to accommodate four new routines. The analysis of the updated data (350 mutations) gives the following informations: (i) 63% of the mutations are missense, and only 20% occur in CpG dinucleotides; (ii) although the mutations are widely distributed throughout the gene, there is an excess of mutations in exons 4 and 9, and a deficit in exons 13 and 15; (iii) the analysis of the distribution of mutations located within the ligand-binding domain shows that 74% of the mutations in this domain affect a conserved amino-acid, and that they are mostly confined in the C-terminal region of the repeats. Conversely, the same analysis in the EGF-like domain shows that 64% of the mutations in this domain affect a non-conserved amino-acid, and, that they are mostly confined in the N-terminal half of the repeats. The database is now accessible on the World Wide Web at http://www.umd.necker.fr

THE LDL RECEPTOR AND HYPERCHOLESTEROLEMIA

The LDL receptor is a 160 kDa transmembrane glycoprotein ubiquitously distributed, playing a major role in cholesterol homeostasis (1). Impairement of LDL receptor activity results in the accumulation of LDL cholesterol in the circulation leading to familial hypercholesterolemia (FH). Affected individuals display arcus corneae, tendon xanthomas and premature symptomatic coronary heart disease (2). FH is an autosomal dominant disease, homozygotes being more severely affected than heterozygotes. FH is also one of the most common inherited disorders with frequencies of heterozygotes and homozygotes estimated to be 1/500 and $1/10^6$, respectively. In certain communities FH frequency is higher due to founder effects (3). The LDL receptor gene (LDLR) lies on the short arm of chromosome 19 (19p13.1–13.3) (4,5). It contains 18 exons encoding the six functional domains of the mature protein: Signal peptide, ligand-binding domain, epidermal growth factor (EGF) precursor like, O-linked sugar, transmenbrane and cytoplasmic (6). To date,

*To whom correspondence should be addressed at: INSERM U383, Clinique Maurice Lamy, Hôpital Necker-Enfants Malades, 149-161 rue de Sèvres, 75743 Paris Cedex 15, France. Tel: +33 1 44 49 44 85; Fax: +33 1 47 83 32 06; Email: boileau@ceylan.necker.fr

444 mutations in the LDLR gene have been identified that are distributed as follows: 350 point mutations (77%), 68 major rearrangements (15%), 20 splice mutations (4%), 6 mutations in the promoter sequence (1%) (3,7).

THE LDLR DATABASE

This second version of the LDLR database contains 350 entries. Table 1 shows the 140 new entries of the database corresponding to mutations either recently published or contributed by the co-authors of this paper (8–31). It is not intended to replace primary publications, although it does contain unpublished data. As in the previous edition, mutation names are given according to Beaudet *et al.* (32) and are often followed by the name of the city or country from which the proband's family originated. For each mutation, information is provided at several levels: gene (exon and codon number, wild type and mutant codon, mutational event, mutation name), protein (wild type and mutant amino acid, affected domain, activity, mutation class), personal (ethnic background, age, sex, body mass index, familial history of coronary heart disease), clinical (values of plasma total cholesterol, LDL-cholesterol, HDL-cholesterol and triglycerids, presence or absence of xanthomas, arcus corneae and symptomatic coronary heart disease) and impact (private, recurrent, founder). We have included possible recurrent mutations (when no comparable haplotypes of the LDLR gene where available) in two instances: (i) when carriers of the same mutation were from distant ethnic or geographic background, and if not (ii) when clinical data were provided for the mutations to allow analysis of phenotypic variability. This last point concerns mutations W23X identified in probands of German–Canadian and German origin, 533ins8 and R395Q identified in probands from Germany, D200G identified in probands of Afrikaner and British origin, S285L identified in probands of Afrikaner and Dutch origin and P664L identified in probands of Belgian, Flemish–Walloon and Dutch origin. The ambiguity between recurrent and founder mutations will only be solved when a consensus will be reached on the polymorphic sites of the LDLR gene that should be systematically typed. Finally, since many teams now systematically screen the whole gene, two-mutations alleles are now being reported. Eleven of these appear in Table 2 (18,33–37). They are not included in the mutations file of the database since it cannot, at present, accomodate two mutations on a single allele.

NEWLY DEVELOPED SOFTWARE ROUTINES

The software package contains routines for the analysis of the LDLR database that were developed with the 4th dimension^{R} (4D) package from ABI. The purpose of the software is to facilitate the mutational analysis of the LDLR gene at the molecular level and to provide the tools to promote the analysis of relationships between phenotype and genotype. Initially, six specific routines were developed (3). Four new routines have been added to the software: (i) «Restriction enzyme» appears on the first page of the mutation record. If the mutation modifies a restriction site, the program shows a restriction map displaying the new or abolished site and the enzymes of interest (Table 1, Column I). (ii) «Amino acid type search» studies the mutations with respect to phylogenic conservation. In effect, the LDLR gene has been identified, sequenced and converted to protein sequence in four mammalian species [complete coding sequence of the chinese hamster (SWISS-PROT accession number: p35950), the rabbit (p20063), the rat (p35952) and the mouse (p35951) LDL receptor] and in the xenope (38). The identity at the amino acid level between the human and chinese hamster, rabbit, rat, mouse and xenopus sequences are 81%, 79%, 77%, 76% and 70%, respectively. Therefore, the routine lists the mutations affecting conserved or non-conserved amino acids in the four mammals, in the xenope, or in all these sequences. (iii) «Phylogeny» studies the distibution of mutations (missense, stop and frameshift) in conserved amino acids between humans and mammals or vetebrates and in amino acids specifically found in the human protein. (iv) «Binary comparison» compares two mutation groups, each group being defined by distinct research criteria chosen from the database records (molecular, clinical, personal, etc.). The result can be displayed as either of several graphic representations (by amino acids, by exon, or by protein domain) of the distribution of the sorted mutations. Furthermore, the sorted mutations can also appear in a cumulated or detailed format (insertion, deletion, missense, nonsense).

RESULTS OF THE FIRST MOLECULAR ANALYSIS

The results of the first molecular analysis of the 350 point mutations of the database shows that 63% of the mutations are missense, and only 20% occur in CpG dinucleotides in opposition to the 32% observed in other human disease genes (39). The origin of this deficit is unknown. Although the mutations are widely distributed throughout the gene, there is an excess of mutations in exon 4 (P = 0.001) coding for the three central repeats of the ligand binding domain, and in exon 9 (P = 0.01) coding for the NH_2 end of the central region of the EGF precursor like domain, between repeats B and C. Conversely, there is a deficit of mutations in exon 13 (P = 0.001) coding for the COOH end of the central region of the EGF precursor like domain, between repeats B and C, and in exon 15 (P = 0.001) coding for the O-linked sugar domain. These mutation hot- or cold-spots cannot be attributed to a technological bias since most teams screened the 18 exons of the LDLR gene. The analysis of the distribution of mutations in the ligand-binding domain, after alignment of the seven repeats, shows that 74% of the mutations in this domain affect a conserved amino acid, and that they are mostly located in the C-terminal region of the repeats. Conversely, the same analysis in the EGF-like domain, after alignment of the three repeats, shows that 64% of the mutations in this domain affect a non-conserved amino acid, and that they are mostly clustered in the N-terminal half of the repeats. Finally, the investigation of genotype/phenotype correlations remains difficult since clinical data are usually incomplete in many published mutation reports. Furthermore, many mutations were identified in compound heterozygotes and the clinical data provided results from the combined effect of the two mutations. To overcome this shortage, we are currently developing an entry in the Web site that will facilitate the input of high quality clinical information for each mutation.

DATABASE ON THE WEB

The LDLR database is now accessible through the World Wide Web at http://www.umd.necker.fr . Users of the database must cite this article. Finally, notification of omissions and errors in the

Table 1. The 140 new mutation reports of the LDLR database

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
213	1	1	-21	ATG	TTG	A->T	No	Nla III -	M-21L	Met	Leu	SP	1	Htz	Wa		P	S. Afr. Indian	26
356	2	91	10	GAG	TAG	G->T	No	Rma I +	E10X	Glu	Stop	LB1	1	Htz	Wa		?	Flemish - Walloon	28
248	2	91	10	GAG	TAG	G->T	No	Spe I +	E10X	Glu	Stop	LB1	1	Htz	Wa		? - F 6/30	Spanish	13
335	2	91	10	GAG	TAG	G->T	No	Spe I +	E10X - MOROCCO	Glu	Stop	LB1	1		Wa		?	Jewish Ashkenazi	23
215	2	131	23	TGG	TAG	G->A	No	BsiY I -	W23X	Trp	Stop	LB1	1	Htz	Wa		? - F	Danish	27
271	2	131	23	TGG	TAG	G->A	No	BsiY I -	W23X	Trp	Stop	LB1	1	Htz	Wa		? - F	German	*4
300	2	131	23	TGG	TAG	G->A	No	BsiY II -	W23X	Trp	Stop	LB1	1	Htz	Wa		?	Belgian	**2
316	2	139	26	GAT	AAT	G->A	Yes		D26N - HYOGO	Asp	Asn	LB1		Htz	Wa		P	Japanese	*3
261	2	148	29	GCT	ACT	G->A	Yes	Dra III +	A29T	Ala	Thr	LB1		Htz	Wa		?	Austrian	*2
283	2	148	29	GCT	ACT	G->A	Yes	Dra III +	A29T	Ala	Thr	LB1		Htz	Wa		?	S. Afr. Coloured	**1
324	2	166	35	TCT	CCT	T->C	No	Hae III +	S35P	Ser	Pro	LB1			Wa		F 2/742	Norwegian	21
222	3	232	57	CGT	TGT	C->T	Yes	Hae III -	R57C	Arg	Cys	LB2					P	British	8
242	3	232	57	CGT	TGT	C->T	Yes	Hae III -	R57C	Arg	Cys	LB2					P	S. Afr. Indian	12
282	3	253	64	CAG	TAG	C->T	No		Q64X	Gln	Stop	LB2	1	Htz	Wa		?	S. Afr. Jewish	**3
216	3	259	66	TGG	GGG	T->G	No	BsiY I +	W66G	Trp	Gly	LB2		Htz	Wa		F	Danish	27
277	3	259	66	TGG	GGG	T->G	No	BsiY I +	W66G	Trp	Gly	LB2	3 or 5	Htz	Wa		?	German	*4
286	3	259	66	TGG	GGG	T->G	No	BsiY I +	W66G	Trp	Gly	LB2		Hmz	ab	?	?	Afrikaner	*1
223	3	260	66	TGG	TAG	G->A	No	Xba I +	W66X	Trp	Stop	LB2					P	British	8
224	3	266	68	TGC	TAC	G->A	No	Rsa I +	C68Y	Cys	Tyr	LB2					P	British	8
225	3	268	69	GAT	AAT	G->A	Yes		D69N	Asp	Asn	LB2					P	British	8
263	3	283	74	TGC	GGC	T->G	No		C74G	Cys	Gly	LB2		Htz	Wa		P	Spanish	18
227	3	301	80	GAG	del1a	Stop at 204			301delG	Glu	Fr.	LB2					F 3/791	British	8
262	3	311	83	TGT	TTT	G->T	No		C83F	Cys	Phe	LB2		Htz	Wa		P	British	*1
240	3-4	313	84	CCC	TCC	C->T	No		P84S	Pro	Ser	LB3		Htz	Wa		P	Finnish	11
217	4	335	91	GAC	del10b	Stop at 201			335del10	Asp	Fr.	LB3		Htz	Wa		P	Danish	9
355	4	337	92	GAG	TAG	G->T	No	Rma I +	E92X	Glu	Stop	LB3	1	Htz	Wa		?	Flemish	28
317	4	344	94	CGC	CAC	G->A	Yes		R94H - FUKUOKA	Arg	His	LB3		Htz	Wa		P	Japanese	*3
301	4	347	95	TGC	TTC	G->T	No	Fnu4H I -	C95F	Cys	Phe	LB3		Htz	Wa		P	Belgian	**2
252	4	429	122	TGC	TGA	C->A	No	Age I +	C122X	Cys	Stop	LB3	1	Htz	Wa		F 13/70 (Wa.) 3/80 (Fl.)	Flemish - Walloon	15
269	4	460	133	CAG	TAG	C->T	No	Rma I +	Q133X	Gln	Stop	LB4		Htz	Wa		P	Spanish	18
325	4	465	134	TGC	TGA	C->A	No		C134X	Cys	Stop	LB4			Wa		F 12/742	Norwegian	21
228	4	479	139	TGC	TAC	G->A	No		C139Y	Cys	Tyr	LB4					P	British	8
299	4	482	140	ATC	ACC	T->C	No		I140T	Ile	Thr	LB4		Htz	Wa		P	Costa Rican	*1
276	·4	500	146	TGC	TAC	G->A	No		C146Y	Cys	Tyr	LB4		Htz	Wa		P	German	*4
229	4	502	147	GAC	AAC	G->A	Yes		D147N	Asp	Asn	LB4					P	British	8
249	4	518	152	TGC	del1b	Stop at 204		Taq I +	518delG	Cys	Fr.	LB4	1	Htz	Wa		F 4/30	Spanish	13
272	4	518	152	TGC	TAC	G->A	No		C152Y	Cys	Tyr	LB4		Htz	Wa		P	German	*4
344	4	519	152	TGC	TGG	C->G	No	BsiY I +	C152W	Cys	Trp	LB4		Htz	Wa		P	German	*5
326	4	523	154	GAT	TAT	G->T	No		D154Y	Asp	Tyr	LB4			Wa		P	Norwegian	21
302	4	527	155	GGC	GTC	G->T	No	BspW I -	G155V	Gly	Val	LB4		Htz	Wa		P	Belgian	**2
264	4	530	156	TCG	TTG	C->T	Yes		S156L	Ser	Leu	LB4		Htz	Wa		R	Spanish	18
345	4	533	157	GAT	ins8b	Stop at 178			533ins8	Asp	Fr.	LB4		Htz	Wa		? - F	German	*5
343	4	554	164	AGG	ins1b	Stop at 178			553insG	Arg	Fr.			Htz	Wa		P	German	*5
338	4	556	165	GGT	del1a	Stop at 204			556delG - ISRAEL	Gly	Fr.				ab	65	P	Arab Moslem	23
290	4	558	165	GGT	ins1c	Stop at 178			558insG	Gly	Fr.			Htz	Wa		P	German	19
295	4	617	185	AGT	del1b	Stop at 204			617delG	Ser	Fr.	LB5		Htz	Wa		P	French	**3
230	4	646	195	TGT	del1a	Stop at 204			646delT	Cys	Fr.	LB5					P	British	8
342	4	661	200	GAC	AAC	G->A	Yes		D200N	Asp	Asn	LB5		Htz	Wa		P	German	*5
266	4	662	200	GAC	GGC	A->G	No	Msp I +	D200G	Asp	Gly	LB5	2B	Htz	Wa		?	Spanish	18
273	4	662	200	GAC	GGC	A->G	No	Msp I +	D200G	Asp	Gly	LB5	2B	Htz	Wa		?	German	*4
280	4	662	200	GAC	GGC	A->G	No	Msp I +	D200G	Asp	Gly	LB5		Htz	Wa		? - F	Afrikaner	*1
243	4	661	200	GAC	TAC	G->T	No		D200Y	Asp	Tyr	LB5		Htz	Wa		?	S. Afr. Indian	12
265	4	661	200	GAC	TAC	G->T	No		D200Y	Asp	Tyr	LB5		Htz	Wa		?	Spanish	18
285	4	671	203	GAC	GCC	A->C	No	Hae III +	D203A	Asp	Ala	LB5		Htz	Wa		P	S. Afr. Coloured	**1
250	4	681	206	GAC	ins7c	Stop at 216			681ins7	Asp	Fr.	LB5	1	Htz	Wa		P	French-Canadian	14
239	4	681	206	GAC	ins8c	ins			681ins8	Asp	Fr.	LB5		Htz	Wa		P	Costa Rican	10
245	4	682	207	GAG	AAG	G->A	Yes		E207K	Glu	Lys	LB5	2B	Htz	Wa		? - F 2/7	S. Afr. Indian	12
336	4	682	207	GAG	CAG	G->C	No	EcoR II +	E207Q - IRAQ	Glu	Gln	LB5	2B		Wa		?	Jewish Ashkenazi	23
354	5	731	223	TCT	TAT	C->A	No		S223Y	Ser	Tyr	LB6		Htz	Wa		P	Flemish	28
231	5	736	225	GGA	del1a	Stop at 263			736delG	Gly	Fr.	LB6					P	British	8
255	5	757	232	CGG	TGG	C->T	Yes	Msp I -	R232W	Arg	Trp	LB6		Htz	ab	2140+5: G>A	P	Austrian	*2
358	5	767	235	GAC	GAG	A->G	No		D235E	Asp	Glu	LB6		Htz	Wa		?	Finnish	25
256	6	828	255	TGC	TGA	C->A	No		C255X	Cys	Stop	LB7		Htz	Wa		F 2/508	Austrian	*2
353	6	829	256	GAG	AAG	G->A	Yes		E256K	Glu	Lys	LB7		Htz	Wa		?	Walloon	28
303	6	855	264	CAC	CAA	C->A	No		H264Q	His	Gln	LB7		Htz	Wa		P	Belgian	**2
341	6	862	267	GAA	AAA	G->A	Yes		E267K	Glu	Lys	LB7		Htz	Wa		P	German	*5
253	6	902	280	GAC	GGC	A->G	No		D280G	Asp	Gly	LB7		Htz			F 5/115	Greek	16
294	6	910	283	GAC	AAC	G->A	No	BstK I -	D283N	Asp	Asn	LB7	2B		Wa		?	Irish	**3
278	6	917	285	TCA	TTA	C->T	No		S285L	Ser	Leu	LB7	2B	Hmz	aa		?	German	29
279	6	917	285	TCA	TTA	C->T	No		S285L	Ser	Leu	LB7	2B	Htz	Wa		? - F	Afrikaner	*1
232	6	931	290	AAA	del2a	Stop at 298			931delAA	Lys	Fr.	LB7					P	British	8
233	6	938	292	TGC	TAC	G->A	No	Rsa I +	C292Y	Cys	Tyr	LB7					F 4/791	British	8
234	7	979	306	CAC	TAC	C->T	No		H306Y	His	Tyr	EGFA					P	British	8
331	7	985	308	TGC	GGC	T->G	No		C308G - POLAND	Cys	Gly	EGFA			Wa		P	Jewish Ashkenazi	23
318	7	1012	317	TGC	CGC	T->C	No	Hae II +	C317R - GIFU	Cys	Arg	EGFA		Htz	Wa		P	Japanese	*3
235	7	1024	321	GAC	AAC	G->A	Yes		D321N	Asp	Asn	EGFA					P	British	8
304	7	1027	322	GGC	AGC	G->A	Yes	Alu I -	G322S	Gly	Ser	EGFA	2Bor5	Htz	Wa		?	Belgian	**2
337	7	1027	322	GGC	AGC	G->A	Yes	Alu I -	G322S - SYRIA	Gly	Ser	EGFA	2Bor5		Wa		?	Jewish Ashkenazi	23
236	7	1033	324	CAG	TAG	C->T	No	Pvu II -	Q324X	Gln	Stop	EGFA					P	British	8
333	7	1046	328	CAG	del1b	Stop at 368			1046delA - RUSSIA/HUNGARY	Gln	Fr.	EGFA			Wa		P	Jewish Ashkenazi	23
254	7	1048	329	CGA	TGA	C->T	Yes		R329X	Arg	Stop	EGFA					? - F 9/78	British	17
292	7	1048	329	CGA	TGA	C->T	Yes		R329X	Arg	Stop	EGFA		Htz	Wa		? - F	German	19
305	7	1056	331	TGC	TGG	C->G	No		C331W	Cys	Trp	EGFA		Htz	Wa		P	Belgian	**2
319	8	1066	335	GAT	CAT	G->C	No	Taq I -	D335H - KANAGAWA	Asp	His	EGFB		Htz	Wa		P	Japanese	*3
306	8	1066	335	GAT	TAT	G->T	No	Taq I -	D335Y	Asp	Tyr	EGFB		Htz	Wa		?	Belgian	**2
327	·8	1066	335	GAT	TAT	G->T	No	Taq I -	D335Y	Asp	Tyr	EGFB			ab	?	? - F 8/742	Norwegian	21
251	8	1124	354	TAC	TGC	A->G	No	Fnu4H I +	Y354C	Tyr	Cys	EGFB			Wa		F 2/88	French-Canadian	14
296	8	1133	357	CAG	CCG	A->C	No	Bsp1286 I +	Q357P	Gln	Pro	EGFB		Htz	Wa		P	British	**3
284	8	1154	364	CTG	CGG	T->G	No	Alu I -	L364R	Leu	Arg	EGFB		Htz	Wa		P	S. Afr. Coloured	**1
214	8	1176	371	TGC	TGA	C->A	No		C371X	Cys	Stop	EGFB	1	Htz	Wa		P	S. Afr. Indian	26
328	9	1201	380	CTC	GTC	C->G	No	Mae II +	L380V	Leu	Val	EGF			Wa		F 5/742	Norwegian	21
320	9	1207	382	TTC	CTC	T->C	No		F382L - WAKAYAMA	Phe	Leu	EGF		Htz	Wa		P	Japanese	*3
218	9	1210	383	ACC	CCC	A->C	No		T383P	Thr	Pro	EGF		Htz	Wa		P	Danish	27
246	9	1215	384	AAC	AAG	C->G	No	Msp I -	N384K	Asn	Lys	EGF		Htz	Wa		P	S. Afr. Indian	12
314	9	1217	385	CGG	CAG	G->A	Yes	Msp I -	R385Q	Arg	Gln	EGF		Htz	Wa		P	S. Afr. Black - Sotho	**2
332	9	1216	385	CGG	TGG	C->T	Yes	Msp I -	R385W - RUMANIA	Arg	Trp	EGF			Wa		P	Jewish Ashkenazi	23
307	9	1243	394	GAC	CAC	G->C	No	Fnu4H I +	D394H	Asp	His	EGF		Htz	Wa		P	Belgian	**2
291	9	1247	395	CGG	CAG	G->A	Yes	Msp I -	R395Q	Arg	Gln	EGF		Htz	Wa		?	German	19
339	9	1247	395	CGG	CAG	G->A	Yes	Msp I -	R395Q	Arg	Gln	EGF		Htz	Wa		?	German	*6
237	9	1246	395	CGG	TGG	C->T	Yes	Msp I -	R395W	Arg	Trp	EGF					? - F4/791	British	8
329	9	1246	395	CGG	TGG	C->T	Yes	Msp I -	R395W	Arg	Trp	EGF			Wa		? - F 9/742	Norwegian	21
308	9	1257	398	TAC	TAG	C->G	No	Rsa I -	Y398X	Tyr	Stop	EGF		Htz	Wa		P	Belgian	**2
288	9	1284	407	AAC	AAG	C->G	No	Mae II -	N407K	Asn	Lys	EGF		Hmz	ab	6	P	S. Afr. Jewish	**3
212	9	1285	408	GTG	ATG	G->A	Yes	Mae II -	V408M	Val	Met	EGF	5	Hmz	aa		F	Greek	30

Table 1. *continued*

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
321	9	1285	408	GTG	ATG	G->A	Yes	Mae II -	V408M - OSAKA	Val	Met	EGF	5	Htz	Wa		?	Japanese	*3
347	9	1294	411	CTG	GTG	C->G	No		L411V	Leu	Val	EGF		Htz	Wa		P	German	*5
267	9	1301	413	ACG	AAG	C->A	No		T413K	Thr	Lys	EGF	5	Htz	Wa		?	Spanish	18
309	9	1301	413	ACG	AGG	C->G	No		T413R	Thr	Arg	EGF		Htz	Wa		? -F	Belgian	**2
352	9	1301	413	ACG	AGG	C->G	No		T413R	Thr	Arg	EGF		Htz	Wa		? -F	Flemish - Walloon	28
330	9	1307	415	GTG	GCG	T->C	No	Msc I -	V415A	Val	Ala	EGF			Wa		P	Dutch	22
298	9	1329	422	TGG	TGC	G->C	No		W422C	Trp	Cys	EGF	2Bor5	Htz	Wa		?	S. Afr. English - British	**3
258	10	1372	437	AGA	del2a	Stop at 438			1372delAG	Arg	Fr.	EGF		Htz	Wa		P	Austrian	*2
323	10	1469	469	TGG	TAG	G->A	No	Rma I +	W469X	Trp	Stop	EGF	1	Hmz	aa		P	German	29
257	10	1474	471	GAC	AAC	G->A	Yes	Xcm I +	D471N	Asp	Asn	EGF		Htz	Wa		F 2/494	Austrian	*2
334	10	1567	502	GTG	del9a	del		Bam HI -	1567del9 - IRAQ	Val	Fr.	EGF			Wa		P	Jewish Ashkenazi	23
289	11	1646	528	GGT	GAT	G->A	No		G528D	Gly	Asp	EGF	2A	Htz	Wa		?	Greek	**3
219	11	1650	529	GTG	del1c	Stop at 546			1650delG	Val	Fr.	EGF		Htz	Wa		P	Danish	27
220	12	1730	556	TGG	TCG	G->C	No		W556S	Trp	Ser	EGF	1	Htz	Wa		F 8/61	Danish	9
351	12	1775	571	GGG	GAG	G->A	No		G571E	Gly	Glu	EGF	5	Htz	Wa		?	Flemish - Walloon	28
322	12	1784	574	CGG	CAG	G->A	Yes	Msp I -	R574Q	Arg	Gln	EGF		Htz	Wa		P	Japanese	*3
310	12	1823	587	CCC	CTC	C->T	No		P587L	Pro	Leu	EGF		Htz	Wa		P	Belgian	**2
311	12	1840	593	TTT	del2a	Stop at 600			1840delTT	Phe	Fr.	EGF		Htz	Wa		P	Belgian	**2
350	13	1864	601	GAT	TAT	G->T	No	EcoR V -	D601Y	Asp	Tyr	EGF		Htz	Wa		P	Flemish	28
349	13	1978	639	CAG	TAG	C->T	No	Rma I +	Q639X	Gln	Stop	EGF		Htz	Wa		P	Flemish	28
259	14	1998	645	TGG	TGA	G->A	No		W645X	Trp	Stop	EGFC		Htz	Wa		P	Austrian	*2
268	14	2000	646	TGT	TAT	G->A	No		C646Y	Cys	Tyr	EGFC	2A	Htz	Wa		R	Spanish	18
346	14	2001	646	TGT	TGA	T->A	No		C646X	Cys	Stop	EGFC			Wa		P	Swedish	31
357	14	2054	664	CCG	CTG	C->T	Yes	Pst I +	P664L	Pro	Leu	EGFC					? - F 7/915	Dutch	24
312	14	2054	664	CCG	CTG	C->T	Yes	Pst I +	P664L	Pro	Leu	EGFC	2B	Htz	Wa		? -F	Belgian	**2
348	14	2054	664	CCG	CTG	C->T	Yes	Pst I +	P664L	Pro	Leu	EGFC	2B	Htz	Wa		? -F	Flemish - Walloon	28
315	14	2056	665	CAG	TAG	C->T	No	**Aci I -**	Q665X	Gln	Stop	EGFC		Htz	Wa		P	Costa Rican	20
270	14	2085	674	ACC	del19c	Stop at 701			2085del19	Thr	Fr.	EGFC		Htz	Wa		P	Spanish	18
293	14	2092	677	TGC	del1a	Stop at 707			2092delT	Cys	Fr.	EGFC		Htz	Wa		?	Greek-French	**3
260	14	2093	677	TGC	TAC	G->A	No		C677Y	Cys	Tyr	EGFC		Htz	Wa		F 3/530	Austrian	*2
313	14	2096	678	CCG	CTG	C->T	Yes	Msp I -	P678L	Pro	Leu	EGFC		Htz	Wa		? -F	Belgian	**2
221	15	2177	705	ACC	ATC	C->T	No		T705I	Thr	Ile	OLS		Htz	Wa		?	Danish	27
244	16	2356	765	AGC	TGC	A->T	No		S765C	Ser	Cys	OLS		Hmz	aa		P	S. Afr. Indian	12
281	16-17	2389	776	GTG	ATG	G->A	Yes	Nla III +	V776M	Val	Met	TM		Htz	Wa		?	Afrikaner	*1
340	17	2392	777	CTC	del9a	del			2392del9	Leu	Fr.	TM		Htz	ab	186	P	German	*6
287	17	2441	793	CGG	CAG	G->A	Yes	Alu I +	R793Q	Arg	Gln	CP		Htz	Wa		P	S. Afr. Black - Xhosa	**2

Each line represents a single LDLR mutation report. The columns contain the following informations and abbreviations:

A: Report number.

B: Exon number in which the mutation occured. Exons are numbered according to Südhof *et al*. (6) with respect to the translational initiation site given by Yamamoto *et al*. (5).

C: Nucleotide position in which the mutation occured.

D: Codon number in which the mutation occured. Codons are numbered according to Yamamoto *et al*. (5). Therefore, the 21 amino acids of the signal peptide (exon 1) are numbered in negative (from –21 to –1). Codon number 1 is the last codon of exon 1 and encodes the first amino acid (Ala) of the mature LDL receptor. If the mutation spans more than one codon, e.g., there is a deletion of several bases, only the first (5′) deleted codon is entered.

E: Normal base sequence of the codon in which the mutation occured.

F: Mutated base sequence of the codon in which the mutation occured. If the mutation is a base pair deletion or insertion, this is indicated by «del» or «ins» followed by the number of bases deleted or inserted and the position of this deletion or insertion in the codon (a, b or c). The nucleotide position is the first that is deleted or the one preceding the insertion. For example, «del19c» is a deletion of 19 bases including the third base of the codon, «ins8b» is an insertion of 8 bases occuring between the second and the third base of the codon.

G: Concerns base substitutions. It gives the base change, by convention, read from the coding strand. If the mutation predicts a premature protein-termination, the novel stop codon position is given, e.g., «stop at 204».

H: Concerns events occuring at a CpG dinucleotide (only C→T or G→A).

I: Concerns the restriction site that is lost, e.g., «Msp I –», or created, e.g., «Taq I +», by the mutation.

J: Mutation name according to Beaudet *et al*. (32). Missense mutations are designated by the codon number flanked by the single letter code of the normal amino acid prior and of the mutant amino acid after (e.g., Val to Met at codon 408 is designated «V408M»). Nonsense mutations are designated similarly exept that X is used to indicate any termination codon (e.g., Cys to stop at codon 134 is designated «C134X»). Frameshift, insertion and deletion mutations are designated by the nucleotide number followed by «ins» for insertion or «del» for deletion. The nucleotide position is the first that is deleted or the one preceding it in the case of insertions. Exact nucleotides are indicated for two or less bases (e.g., 617delG). For three or more bases, the insertion or deletion is specified by the size of the change (e.g. 681ins8 indicates a 8 bp insertion starting after nucleotide 681). For many of the mutations that have been reported this nomenclature has not been used. Therefore, the original name also appears in this column. These names were given according to the population or the city in which the mutation was reported first (e.g. TOKYO).

K: Wild type amino acid.

L: Mutant amino acid. Deletion and insertion mutations which result in a frameshift are designated by «Fr. »; Nonsense mutations are designated by «Stop».

M: Protein domain in which the mutation occurs. «SP» for the signal peptide, «LB» for the ligand binding domain, «EGF» for the Epidermal Growth Factor precursor like domain, «OLS» for the O-linked sugar chains domain, «TM» for the transmembrane domain, and «CP» for the cytoplasmic domain. In the ligand-binding domain (LB), each of the seven repeats are numbered separately and according to their position with respect to the N-terminal end of the protein.

N: Functional class as defined by Hobbs *et al*. (40).

O: Clinical status according to Goldstein *et al*. (2): «Hmz» indicates homozygotes and «Htz» indicates heterozygotes.

P: Genotype: «aa» indicates homozygotes, «ab» indicates compound heterozygotes, and «Wa» indicates heterozygotes. Empty cases appear when no information is available.

Q: Number of the report in which the second mutation identified in a compound heterozygote is described. When the second mutation is one of those omitted in the database, this mutation is briefly described with respect to the coding sequence. Finally, «?» indicates that the second mutation has not been identified.

R: Recurence of the mutation. «F» indicates a founder effect, «F 2/140» indicates that the mutation was found in two unrelated probands in a sample 140 FH patients, «R» indicates recurrent mutations, «?» indicates mutations that have been identified in at least two unrelated probands of different ethnic backgrounds but for which LDLR gene haplotypes are not described, «?–F» indicates mutations for which LDLR gene haplotypes are not described (or incomplete) and that either are associated with a founder effect in the proband's ethnic or geographic origin, or have been identified in at least two unrelated probands of the same ethnic or geographic background, and «P» indicates mutations identified, to date, in a single proband.

S: Ethnic or geographic background of the proband.

T: Reference number indicating the publication in which the mutation is described. Full citations (authors, year, title, journal, volume, pages) are provided with the database. If the same mutation has been reported for the same patient in different papers, only one entry is made.

*Indicates the co-authors who provided the information: *1 (Rochelle Thiart and Maritha J. Kotze), *2 (Helena Schmidt and Gert M. Kostner), *3 (Yasuko Miyake and Taku Yamamura), *4 (Heike Baron and Herbert Schuster), *5 (Margit Ebhardt and Manfred Stuhrmann) and *6 (Hartmut Schmidt).

**Indicates submitted papers: **1 (O.Loubser *et al*.), **2 (A.Peeters *et al*.) and **3 (M.Callis *et al*.).

Table 2. Each line represents a single LDLR mutation report

First mutation											Second mutation											O	R	S	T
J	B	C	D	E	F	G	K	L	M	P	J	B	C	D	E	F	G	K	L	M	P				
W-18X	1	12	-18	TGG	TGA	G->A	Try	Stop	SP	Wa	E256K	6	829	256	GAG	AAG	G->A	Glu	Lys	LB7	Wa	Htz	P	Spanish	18
Q71E	3	274	71	CAA	GAA	C->G	Gln	Glu	LB2	Wa	313+1(G->C)	3	313+1	-	-	-	G->C	-	-	LB3	Wa	Htz	P	Spanish	18
C95R	4	346	95	TGC	CGC	T->C	Cys	Arg	LB3	Wa	D679E	14	2100	679	GAC	GAG	C->G	Asp	Glu	EGFC	Wa	Htz	P	Spanish	18
654ins6	4	654	197	GGT	ins6c	ins	Gly	Fr	LB5		657del5	4	657	198	GGC	del5c	del	Gly	Fr	LB5			P	German	37
C281Y	6	905	281	TGC	TAC	G->A	Cys	Tyr	LB7	Wa	1706-10(G->A)	11	1706-10	-	-	-	G->A	-	-	EGF	Wa	Htz	P	Spanish	18
D333A	8	1061	333	GAT	GTC	A->C	Asp	Ala	EGFB	aa	2140+5(G->A)	14	2140+5	-	-	-	G->A	-	-	EGF	aa	Hmz	P	Austrian	*2
1115del9	8	1115	351	GAG	del9b	del	Glu	Fr	EGFB	Wa	115ins6	8	1115	351	GAG	ins6a	ins	Glu	Fr	EGFB	Wa	Htz	F 2/-	Japanese	34
Q363X	8	1150	363	CAG	TAG	C->T	Gln	Stop	EGFB	Wa	D365E	8	1158	365	GAC	GAG	C->G	Asp	Glu	EGFB	Wa	Htz	P	Cypriot	33
N543H	11	1690	543	AAT	CAT	A->C	Asn	His	EGF	Wa	2393del9	17	2393	777	CTC	del9b	del	Leu	Fr	TM	Wa	Htz	? - F 2/63	Danish	36
N543H	11	1690	543	AAT	CAT	A->C	Asn	His	EGF	Wa	2393del9	17	2393	777	CTC	del9b	del	Leu	Fr	TM	Wa	Htz	? - F 10/184	Dutch	35
A585T	12	1816	585	GCC	ACC	G->A	Ala	Thr	EGF	Wa	G654S	14	2023	654	GGC	AGC	G->A	Gly	Ser	EGFC	Wa	Htz	F 2/530	Austrian	*2

Footnotes as for Table 1.

current version as well as specific phenotypic data would be gratefully received by the corresponding authors.

ACKNOWLEDGEMENTS

This work was supported by grants from GREG (Groupe de Recherche et d'Etude du Génome), Fondation de France, Université René Descartes Paris V, Ministère de l'Education Nationale, de l'Enseignement Supérieur, de la Recherche et de l'Insertion Professionnelle (ACC-SV2), and Faculté de Médecine Necker, France; The South African Medical Research Council and the Universities of Stellenbosch and Free State, South Africa. M.V. is supported by a grant from Ministère de l'Education Nationale, de l'Enseignement Supérieur, de la Recherche et de l'Insertion Professionnelle. Finally, we gratefully acknowledge the help of the many clinicians that collaborated with the co-authors.

REFERENCES

1 Goldstein, J. and Brown, M. (1985) *Annu. Rev. Cell. Biol.*, **1**, 1–39.
2 Goldstein, J. and Brown, M. (1989) In Scriver,C.R., Beaudet,A.L. and Sly,W.S. (Eds) *The Metabolic Basis of Inherited Diseases -6th Edition*. McGraw-Hill, New-York, pp 1215–1250.
3 Varret, M., Rabès, J.-P., Collod-Béroud, G., Junien, C., Boileau, C. and Béroud, C. (1997) *Nucleic Acids Res.*, **25**, 172–180.
4 Lindgren, V., Luskey, K. L., Russell, D. W. and Francke, U. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8567–8571.
5 Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. and Russell, D. W. (1984) *Cell*, **39**, 27–38.
6 Südhof, T. C., Russell, D. W., Goldstein, J. L. and Brown, M. S. (1985) *Science*, **228**, 893–895.
7 Varret, M., Rabès, J.-P. and Boileau, C. (1997) *Med. Sci.*, **13**, In press.
8 Day, I., Whittall, R., O'Dell, S., Haddad, L., Bolla, M., Gudnason, V. and Humphries, S. (1997) *Hum. Mut.*, **10**, 116–127.
9 Jensen, H., Holst, H., Jensen, L., Jorgensen, M., Andreasen, T., Jensen, T., Andreasen, B., Heath, F., Hansen, P., Neve, S., *et al.* (1997) *Atherosclerosis*, **131**, 67–72.
10 Kotze, M. J., Thiart, R., Loubser, O., de Villiers, J. N. P., Santos, M., Vargas, M. A. and Peeters, A. V. (1996) *Hum. Genet.*, **98**, 476–478.
11 Vuorio, A. F., Turtola, H. and Kontula, K. (1997) *Clin. Genet.*, **51**, 191–195.
12 Kotze, M. J., Loubser, O., Thiart, R., de Villiers, J. N. P., Langenhoven, E., Theart, L., Steyn, K., Marais, D. A. and Raal, F. J. (1997) *Clin. Genet.*, **51**, 394–398.
13 Cenarro, A., Jensen, H. K., Civeira, F., Casao, E., Ferrando, J., Gonzalezbonillo, J., Pocovi, M. and Gregersen, N. (1996) *Clin. Genet.*, **49**, 180–185.
14 Assouline, L., Leitersdorf, E., Lambert, M., Reshef, A., Feoli-Fronseca, J. C. and Levy, E. (1997) *Hum. Mut.*, **9**, 555–562.
15 Descamps, O., Hondekjin, J. C., Van Acker, P., Deslypere, J. P. and Heller, F. R. (1997) *Clin. Genet.*, **51**, 303–308.
16 Mavroidis, N., Traeger-Synodinos, J., Kanavakis, E., Drogari, E., Matsaniotis, N., Humphries, S. E., Day, I. N. M. and Kattamis, C. (1997) *Hum. Mut.*, **9**, 274–276.
17 Day, I. N. M., Haddad, L., O'Dell, S. D., Day, L. B., Whittall, R. A. and Humphries, S. E. (1997) *J. Med. Genet.*, **34**, 111–116.
18 Cenarro, A., Jensen, H. K., Casao, E., Civeira, F., Gonzalez-Bonillo, J., Rodriguez-Rey, J. C., Gregersen, N. and Pocovi, M. (1998) *Hum. Mut.*, In press.
19 Thiart, R., Loubser, O., de Villiers, J. P. N., Marx, M. P., Zaire, R., Raal, F. J. and Kotze, M. J. (1998) *Hum. Mut.*, In press.
20 Thiart, R., Loubser, O., de Villiers, J. N. P., Santos, M. and Kotze, M. J. (1998) *Mol. Cell Probes*, In press.
21 Leren, T. P., Tonstad, S., Gundersen, K. E., Bakken, K. S., Rodningen, O. K., Sundvold, H., Ose, L. and Berg, K. (1997) *J. Int. Med.*, **241**, 185–194.
22 Lombardi, P., Defesche, J., Kamerling, S. W. A., Kastelein, J. J. P. and Havekes, L. M. (1997) *Clin. Genet.*, **51**, 286–287.
23 Reshef, A., Nissen, H., Triger, L., Hensen, T. S., Eliav, O., Schurr, D., Safadi, R., Gare, M. and Leitersdof, E. (1996) *Hum. Genet.*, **98**, 581–586.
24 Defesche, J. C., van de Ree, M. A., Kastelein, J. J. P., van Diermen, D. E., Janssens, N. W. E., van Doormaal, J. J. and Hayden, M. R. (1992) *Clin. Genet.*, **42**, 273–280.
25 Koivisto, U. M., Gylling, H., Miettinen, T. A. and Kontula, K. (1997) *Arterioscler. Thromb. Vasc. Biol.*, **17**, 1392–1399.
26 Langenhoven, E., Warnich, L., Thiart, R., Rubinsztein, D. C., van der Westhuysen, D. R., Marais, A. D. and Kotze, M. J. (1996) *Atherosclerosis*, **125**, 111–119.
27 Jensen, H. K., Jensen, L. G., Hasen, P. S., Faergman, O. and Gergersen, N. (1996) *Clin. Chem.*, **42**, 1140–1146.
28 Descamps, O., Hondekjin, J.-C., Vanaker, P., Deslypere, J.-P. and Heller, F.-R. (1997) *Atherosclerosis*, **130** (Supplement), S26.
29 Schuster, H., Keller, C., Wolfram, G. and Zollner, N. (1995) *Aterioscler. Throm. Vasc. Biol.*, **15**, 2176–2180.
30 Schuster, H., Manke , C., Ficher, J., Keller, C., Wolfram, G. and Zollner, N. (1995) *Clin. Genet.*, **48**, 90–92.
31 Ekstrom, U., Abrahamson, M., Sveger, T., Lombardi, P. and Nilsson-Ehle, P. (1995) *Hum. Genet.*, **96**, 147–150.
32 Beaudet, A. L. and Tsui, L.-C. (1993) *Hum. Mut.*, **2**, 245–248.
33 Kotze, M. J., de Villiers, J. N. P., Loubser, O., Thiart, R., Scholtz, C. and Raal, F. J. (1997) *Hum. Genet.*, **100**, 101–103.
34 Yamakawa-Kobayashi, K., Kobayashi, T., Yanagi, H., Shimakura, Y., Satoh, J. and Hamaguchi, H. (1994) *Hum. Genet.*, **93**, 625–628.
35 Lombardi, P., Kamerling, S. W. A., Defesche, J. C., Kastelein, J. J. P. and Havekes, L. M. (1996) *Clin. Genet.*, **50**, 525–526.
36 Jensen, H. K., Jensen, T. G., Faergeman, O., Jensen, L. G., Andresen, B. S., Corydon, M. J., Andreasen, P. H., Hansen, P. S., Heath, F., Bolund, L. and Gregersen, N. (1997) *Hum. Mut.*, **9**, 437–444.
37 Geisel, J., Holzem, G. and Oette, K. (1995) *Hum. Genet.*, **96**, 301–304.
38 Mehta, K. D., Chen, W.-J., Goldstein, J. L. and Brown, M. S. (1991) *J. Biol. Chem.*, **266**, 10406–10414.
39 Cooper, D. N. and Krawczak, M. (1990) *Hum. Genet.*, **85**, 55–74.
40 Hobbs, H. H., Brown, M. S. and Goldstein, J. L. (1992) *Hum. Mutat.*, **1**, 445–466.