Splice-Site Mutations in Atherosclerosis Candidate Genes

Relating Individual Information to Phenotype

Yskert von Kodolitsch, MD; Reed E. Pyeritz, MD, PhD; Peter K. Rogan, PhD

Background—Nucleotide variants in several genes for lipid and methionine metabolism influence the risk of premature atherosclerosis. Ten percent of single nucleotide substitutions in these genes involve mRNA splice sites. The effects of some of these changes on splicing and on phenotypic severity are not inherently obvious.

Methods and Results—Using an information theory-based model, we measured the individual information content (R_i, in bits) of splice sites adjacent to 289 mutations (including 31 splice-site mutations) in the atherosclerosis candidate genes *APOAII*, *APOB*, *APOCII*, *APOE*, *CBS*, *CETP*, *LCAT*, *LIPA*, *LDLR*, and *LPL*. The predictions of information analysis were then corroborated by published mRNA analyses. The R_i values of mutant sites were consistent with either complete (n=17) or partial (n=8) inactivation of these sites. Seven mutations were predicted to activate cryptic splice sites. Predicted inactive mutant sites were associated with either "average" or "severe" dyslipidemia and commensurate reductions in protein levels or activity, whereas mutations expected to exhibit residual splicing had average or "mild" effects on lipid and protein expression.

Conclusions—Information analysis of splice-junction variants in atherosclerosis candidate genes distinguishes inactive from leaky splice sites and identifies activated cryptic sites. Predicted changes in splicing were related to phenotypic severity. **(*Circulation*. 1999;100:693-699.)**

Key Words: atherosclerosis ■ genetics ■ lipids ■ risk factors ■ RNA

Mendelian syndromes that carry a substantial risk for developing occlusive arterial disease frequently alter metabolic pathways involving cholesterol, triglycerides, and sulfurated amino acids. For instance, mutations in the genes coding for cystathione β-synthase (*CBS*), the low density lipoprotein receptor (*LDLR*), the *LDLR* binding region of apolipoprotein B (*APOB)*, apolipoprotein E (*APOE)*, lecithin-cholesterol transferase (*LCAT*), lysosomal acid lipase A (*LIPA*), lipoprotein lipase (*LPL*), and *APOCII* cause occlusive arterial disease. A low risk of atherosclerosis is observed in *APOB* mutations outside the LDL binding region and in cholesteryl ester transfer protein (*CETP*) deficiency, whereas the impact of *APOAII* deficiency on atherosclerosis has not been established[1] (Table 1).

Among the single nucleotide substitutions in these genes, 10% are located within splice junctions (Table 1). The donor GT and acceptor AG dinucleotides at the splice junctions are conserved; however, other sequences at these and other conserved positions are also permissible. Mutations within natural (termed "primary") splice junctions can either completely or partially inactivate these sites. Some nucleotide changes activate splice sites that are not normally recognized ("cryptic" or "secondary" sites). Other splice-site substitutions have no effect on mRNA splicing.[2,3]

Information theory–based models that account for all of the nucleotide variations in splice sites have been used to predict the activities of natural and mutant sites and identify cryptic splice sites.[2] The information content (in bits) of a member of any sequence family describes the degree to which that member contributes to the conservation of the entire family.[4] The effects of all nucleotide changes are detectable, given that information is cumulative over all positions in a splice site.[2]

In previous studies, lipid and protein levels or enzymatic activities of gene products were used to indirectly assess the effects of mutations that predispose a person to atherosclerosis. We used information theory–based models of mRNA splicing to relate the severity of splicing mutations to atherosclerosis phenotypes.

Methods

Selection of Mutations

More than 200 genes have been proposed as influencing the risk of developing atherosclerosis.[1] Of the genes known to alter lipid and methionine metabolism, the Human Gene Mutation Database[5] reveals 10 genes that have splicing mutations (n=31). Missense mutations (n=258) were also analyzed to examine whether these changes might have collateral effects on splicing.[6]

Received December 31, 1998; revision received May 10, 1999; accepted May 25, 1999.

From the Department of Cardiology (Y.v.K.), University Hospital Eppendorf, Hamburg, Germany; Departments of Human Genetics, Medicine, and Pediatrics (R.E.P.), Allegheny General Hospital, Pittsburgh, Pa; and Section of Molecular Genetics and Molecular Medicine (P.K.R.), Children's Mercy Hospital, Kansas City, Mo.

Correspondence to Peter K. Rogan, PhD, Section of Molecular Genetics and Molecular Medicine, Children's Mercy Hospital and Clinics, 2401 Gillham Road, Kansas City, MO 64108. E-mail progan@cmh.edu

Circulation **is available at http://www.circulationaha.org**

TABLE 1. Reference Parameters for Atherosclerosis Candidate Gene Mutations

			Protein (% of Normal)		Plasma Lipid Profile				Reference Plasma Lipid Values*				
Gene	Disorder	Trait	Missense Mutation	Splicing Mutation	TG	TC	LDL-c	HDL-c	TG	TC	LDL-c	HDL-c	Disease Risk
APOAII	Apo A-II deficiency	rec	<1	...	N	N	N	N	129±36.2	231±31	...	57±11	...
APOB	Hypobetalipoproteinemia	cod	0 (31–38)	...	−−	−−	−−	+	49±2 (77±61)	73±23 (127±32)	5±2 (50±18)	61±26 (60±22)	−
	Familial defective apo B-100	dom	...	...	N	+,++	+,++	N	...	...	...	...	+
APOCII	Apo C-II deficiency (type I HLP)	rec	0	0	++	N,+	−−	−−	3337±2570 (257±85)	363±253 (141±84)	...	...	N
APOE	APOE deficiency (type III HLP)	cod	0	...	++	++	N,−	N,−	694±104	440±25	...	37±3	+
CBS	CBS deficiency (homocystinuria)	rec	...	0–40 (7–100)		...	...	...	...	...	...	...	+
CETP	CETP deficiency	dom	0.6 (56)	0 (64)	N,+	+,++	N,−	++	...	...	...	84±25	−
LCAT	LCAT deficiency,	rec	2–43 (44–80)	0–9 (50–57)	N,+	N,+,−	N,+,−	−	...	...	...	...	(+)
	fish eye disease	rec	50 (100)	0–5 (61)	+,++	N,+	N,+	−	(90±31.5)	(178±33)	(134±31)	(31±5.5)	N
LDLR	Familial hypercholesterolemia (type II HLP)	dom	...	<2–30 (35–50)	N,+	++	++	−	101±51 (148±75)	378±170 (368±78)	625±160 (298±78)	34±10 (44±13)	+
LIPA	Cholesterol ester storage disease	rec	...	0–36	N,+	+	N,+	−	208±93	319±78	151±31	43±15	+
LPL	LPL deficiency (type I HLP)	rec	...	0 (20–62)	++	N,+	−−	−−	2398±1595	304±247	19±10	11.3±2.7	(+)

HLP indicates hyperlipoproteinemia; dom, dominant; cod, codominant; rec, recessive; TG, plasma triglycerides; TC, plasma total cholesterol; LDL-c; LDL-cholesterol; HDL-c, HDL-cholesterol; +, increased, and, in some cases, high levels; ++, very high levels; N, normal levels; −, low levels; −−, very low levels; and ..., no information assessed.

*Lipid values are in mg/dL±1 SD for homozygous-affected individuals, and values in parentheses refer to heterozygotes.

Computation of Individual Information

The individual information contents (R_i) of natural and mutant splice sites and coding sequence variants in atherosclerosis candidate genes were compared.[4] The computer programs Scan, MakeWalker, and Lister were used to compute R_i values and display results.[7]

The information contents of potential splice sites were determined for sequences up to 150 nucleotides upstream and downstream of each nucleotide substitution.[2,8] R_i values could not be directly assessed for the *LDLR* and *CBS* genes[2,7] because the terminal nucleotides of the corresponding splice junctions (positions −16 to −25 for acceptor sites, and +6 for donor sites) were unavailable. To model incomplete sites, the possible range of R_i values was defined by computing R_i for the best and worst sequence combinations. These ranges were computed for the missing sequence data from the corresponding coordinates of the donor and acceptor weight matrix elements that produced the maximal and minimal R_i values. Interpretation of variants was possible only if the R_i intervals of the natural and corresponding mutant sites did not overlap. Inactivating mutations were distinguishable from those that resulted in leaky splicing if the R_i interval of the mutant site was consistently either less or greater than 2.4 bits, respectively. Six of 171 variants with incomplete acceptor splice sites could not be interpreted because the R_i intervals of the variant and corresponding natural sequences overlapped.

Criteria for Grading Phenotypic Severity

For each splicing mutation, reference distributions of protein expression, plasma lipids, or plasma homocystine were derived from individuals with other mutations in the corresponding gene from previously published reports (data not shown) (Table 1). Individuals with mutations in the same gene were also distinguished according to genotype because the diagnosis, in some instances, depended on the mode of inheritance. Protein levels and enzymatic and/or binding activity were measured by the same methods for each splice-site mutation and the corresponding reference population; these values are given as the reported range of the percentage of normal protein levels. The mean and 95% confidence intervals (±2SD) of lipid values for each reference population were either derived from published reports or computed from values in published patient cohorts.

TABLE 2. Information Analysis of Splice-Site Variants

Mutation No.	Gene (Accession)	Mutation, Coordinate	Primary Splice Site		Secondary Splice Site		Primary Splice Site Use (% use†)	Phenotype			
			Coordinate	$R_{i,\ natural} \rightarrow R_{i,\ mutant}$*	Coordinate	$R_{i,\ natural} \rightarrow R_{i,\ mutant}$		Genotype‡	Activity (Mass)§	Dyslipidemia‖	Reference
Experimentally verified loss of primary splice site function											
1	*CETP* (M32995)	IVS 10, donor, T→G, 290	289¶	7.4→−0.8	IVS 13, 1254	8.9→8.9	−	12	15	+++	9
2	*LPL* (S71696)	IVS 1, donor, G→C, 14	10	9.7→−3.0			−	11	0 (0)	++	10
3	*LDLR* (L00337)	IVS 3, donor, G→A, 139	Best,# 139 Worst, 139	5.8→−7.0 4.9→−7.9			−	12	Reduced	++	11
4	*LDLR* (L00337)	IVS 3, donor, G→T, 139	Best, 139 Worst, 139	5.8→−2.0 4.9→−2.9			−	12	...	++	12
5	*LDLR* (L00338)	IVS 4, donor, T→C, 398	Best, 397 Worst, 397	5.2→−2.2 4.7→−2.8			−	...	...	...	13
6	*LIPA* (X75494)	IVS 6, acceptor, A→G, 394	395	13.5→5.1			+ (12)	23	Reduced	...	14
7	*LIPA* (X75495)	IVS 8, donor, G→A, 416	417	8.8→5.7			+ (32)	11	4.1	++	15, 16
8	*LPL* (M76722)	IVS 6, acceptor, C→A, 196	198	6.7→2.5			+ (24)	11	2	+	17
Predicted loss of primary splice-site function											
9	*APOA-II* (X04898)	IVS 3, donor, G→A, 1882	1882	9.5→−3.3			−	11	<0.3	++	18
10	*APOB* (M19811)	IVS 5, donor, G→T, 411	411	9.0→1.2			−	12	...	++	19
11	*APOC-II* (M10612)	IVS 2, donor, G→C, 2995	2995	9.0→−0.8			−	11	(0.1)	++	20
12	*LPL* (S71696)	IVS 2, acceptor, G→A, 40	40	7.6→0.1			−	23	0	++	21
13	*LDLR* (L00343)	IVS 9, donor, G→A, 188	Best, 188 Worst, 188	11.3→−1.5 10.4→−2.4			−	12	...	+++	22
14	*LDLR* (L00344)	IVS 10, donor, G→A, 248	Best, 244 Worst, 244	5.8→−16.2 4.8→−17.2			−	...	...	++	23
15	*LDLR* (L00345)	IVS 11, donor, G→T, 135	Best, 135 Worst, 135	8.4→0.6 7.5→−0.3			−	12	...	...	24
16	*LDLR* (L00346)	IVS 12, donor, T→C, 157	Best, 156 Worst, 156	7.3→−16.2 6.3→−15.8			−	12	33	+++	25, 26
17	*LPL* (M76722)	IVS 6, acceptor, C→T, 196	198	6.7→5.1			+ (62)	11	...	...	27
Experimentally verified cryptic splicing											
18	*LPL* (S71696)	IVS 2, donor, G→A, 9	10	9.7→−3.1			−	11	0.3 (0.5)	++	28
19	*LDLR* (L00349)	IVS 15, donor, G→A, 187	Best, 187 Worst, 187	8.2→−4.6 7.3→−5.5	Donor, 66**	5.8→5.8	−	11	11	++	29
20	*APOE* (M10065)	IVS 3, acceptor, A→G, 3779	3780	10.8→5.1	Acceptor, 3726**	8.4→8.4	+ (18)	11	(0.3)	+††	30, 31
Predicted cryptic splicing											
21	*APOB* (M19828)	IVS 24, donor, T→C, 1623	1622	7.9→0.4	Donor, 1632 Donor, 1663	3.4→3.4 6.4→6.4	−	11	(0)	+++	32
22	*CETP* (M32997)	IVS 14, donor, G→A, 1358	1358	7.0→−5.6	Donor, 1362	3.8→3.8	−	11	0	+++	33, 34

TABLE 2. Continued

Mutation No.	Gene (Accession)	Mutation, Coordinate	Primary Splice Site		Secondary Splice Site		Primary Splice Site Use (% use†)	Phenotype			
			Coordinate	$R_{i,\ natural} \rightarrow R_{i,\ mutant}$*	Coordinate	$R_{i,\ natural} \rightarrow R_{i,\ mutant}$		Genotype‡	Activity (Mass)§	Dyslipidemia‖	Reference
Predicted cryptic splicing											
23	*LDLR* (L00342)	Exon 8, donor, G→A, 141	Best, 142 Worst, 142	7.0→4.0 6.5→3.4	Donor, 73 Donor, 89 Donor, 91 Donor, 138	3.4→3.4 3.0→3.0 5.3→5.3 3.7→3.7	+ (33)	12	...	++	35
24	*LDLR* (L00350)	Exon 16, donor, G→A, 93	Best, 94 Worst, 94	7.7→4.6 7.2→4.1	Donor, 63 Acceptor, 93	5.0→5.0 2.9→8.4	+ (32)	12	...	++	36
Predicted unaltered splice-site function											
25	*LCAT* (X04981)	IVS 4, acceptor, T→C, 2327	2348	9.7→9.6			−	12	68	++	37
26	*LDLR* (L00342)	IVS 7, acceptor, T→C, 19	Best, 26 Worst, 26	19.2→19.2 6.4→6.4			−−	23	...	...	38
27	*LDLR* (L00338)	Exon 4, donor, T→G, 393	Best, 397 Worst, 397	5.2→7.3 4.7→6.7			−−	12	...	...	39
28	*LDLR* (L00340)	Exon 6, donor, G→A, 136	Best, 139 Worst, 139	9.1→10.2 8.6→9.6			−−		...	...	40
29	*LDLR* (L00342)	Exon 8, acceptor, A→G, 27	Best, 26 Worst, 26	19.2→20.1 6.4→7.4			−−	23	15–30	...	41
30	*LDLR* (L00342)	Exon 8, acceptor, A→T, 27	Best, 26 Worst, 26	19.2→17.9 6.4→5.2			−−	23	15–30	...	41
31	*LPL* (AF050163)	Exon 4, acceptor, G→A, 950	949	11.1→10.2			−−	11	0	+++	42
32	*LDLR* (L00344)	IVS 9, acceptor, C→T, 6	35**	10.7→10.7	Acceptor, 29 Acceptor, 43	2.6 7.8	+ (100)	12	...	++	43

*$R_{i,\ natural}$ and $R_{i,\ mutant}$ are R_i values of the natural/primary and mutant splice sites.

†Splice-site use predicted to be leaky (+), unaltered (−−), or none (−); for leaky mutations, calculated percent of correctly spliced mRNA is shown (in parentheses).

‡Genotype of mutation designated as homozygous (11), heterozygous (12), or compound heterozygous (23).

§Protein activity and mass are, respectively, the percent normal enzymatic activity/binding and the level of expression.

‖Dyslipidemia is graded as mild (+), average (++), or severe (+++); phenotype not available (...).

¶Cryptic splicing verified in intron 13 (accession M32997, position 1254); R_i=8.9 bits.

#"Best" and "worst" are the maximum and minimum possible R_i values for incomplete sites.

**Additional cryptic sites with lower R_i values are detected.

††Plasma triglyceride is +4.5SD; HDL-cholesterol and total cholesterol are, respectively, +4.5SD and +3.6SD above average.

Stringent criteria for dyslipidemic phenotypes were defined from reference mean values based on populations of individuals with mutations in the same gene. Each lipid parameter (total plasma triglycerides, cholesterol, and LDL and HDL cholesterol) with values within ±2SD of the respective reference population was defined as "average." Individuals with significantly increased lipid levels (>2SD above the mean in the affected reference population) were defined as having "severe" dyslipidemia (eg, Table 2, mutations 1, 13, 16, 22, and 31 [mutation numbers throughout the article are those found in Table 2]). Those carrying mutations that decreased lipid levels >2SD below the average in the reference population (eg, mutation 21) were also designated as severe. Dyslipidemia was defined as "mild" in patients with mutations producing HDL-cholesterol levels >2SD above the average for the reference population (mutations 8 and 20).

Protein levels and/or activities were similarly graded relative to reference values for populations with mutations in the same gene. Protein levels and/or activities were defined as average if they were within the reported range of the percentage of normal for each respective reference population. Reduction of protein levels and/or activities below or above this reported range were graded as severe or mild, respectively.

Clinical severity of atherosclerosis was assessed using age-related expression of angina pectoris or myocardial infarction. The ages of onset of coronary symptoms for 2 or more unrelated patients with identical *LDLR* (mutation 3) and *APOE* (mutation 20) mutations were related to the corresponding age distributions for reference populations with type II (*LDLR*)[44] or type III (*APOE*)[45] familial hyperlipoproteinemia. Clinical signs of atherosclerosis were graded as average for ages of onset within 2SD of the mean age for the reference populations; mild or severe atherosclerosis occurred with average ages of onset above or below the 2SD bound, respectively.

Phenotypic analyses were limited to patients with fasting plasma lipid and homocysteine levels. Because outlier phenotypes were derived using stringent statistical criteria, some patients originally described as severely or mildly affected were categorized as having

TABLE 3. Relationship Between Information Content of Splice-Site Mutations and Phenotype

R_i Value Cutoff, bits	Phenotype*					
	Dyslipidemia			Reduction in Protein Level or Activity		
	Mild	Average	Severe	Mild	Average	Severe
<2.4	0/15	10/15	5/15	1/9	7/9	1/9
≥2.4	2/5	3/5	0/5	2/3	1/3	0/3

*No. of mutations in category/total number.

average phenotypes in the present study (mutations 3 and 7 [reference 15]). Mutations in compound heterozygotes (n=5) were not analyzed because the phenotypic effects of each allele could not be separated.

Criteria for Predicting Splice-Site Function

The effects of nucleotide substitutions were predicted from R_i values on the basis of the following, previously validated criteria[2]:

1. Substitutions resulting in R_i<2.4 bits would completely inactivate primary splice sites and result in severely affected phenotypes.
2. Substitutions with decreased R_i values that were ≥2.4 bits would reduce splicing, thereby producing a milder phenotype. The residual amount of correctly spliced mRNA at leaky sites was computed as the minimum fold change in binding affinity, $2^{\Delta Ri}$ (ΔR_i is the difference between the R_i value of the natural site and that of the variant site). The result was expressed as the maximum percent of normal mRNA.[2,5]
3. Substitutions would activate cryptic splice sites with R_i values exceeding that of the adjacent natural site. Observations of exon skipping or cryptic splicing were used to validate R_i-based predictions whenever these data were available.[2]
4. Nucleotide changes that did not significantly alter R_i would have no effect on splicing.[7] The significance of the change in R_i (ΔR_i) was assessed with a paired *t*-test using a cutoff at 5% significance.[2]

Results

Prediction and Validation of Splicing Defects

We predicted either complete loss of primary splice site function (mutations 1 to 5, 9 to 16, 18, 19, 21, and 22) or leaky splicing (mutations 6 to 8, 17, 20, 23, and 24) on the basis of reductions in R_i values of the corresponding mutant sites. We predicted that mutations 18 to 24 would activate cryptic splice sites. mRNA analyses corroborated these predictions for primary splice sites (mutations 1 to 8 and 18 to 20) and activated cryptic splice sites (mutations 1, 18 to 20, and 32). The R_i value of a polymorphic intron variant within an acceptor site was unchanged (mutation 26), which was expected on the basis of previous studies of polymorphic splice sites.[2,7]

Of the 258 missense mutations analyzed, we only predicted 2 (0.8%) would decrease R_i values and affect splicing at adjacent donor sites (mutations 23 and 24). Five missense substitutions produced insignificant changes in the R_i values in adjacent splice sites (mutations 27 to 31); the remainder had no detectable effect on splicing. Despite the fact that the splice-site sequences for the *LDLR* and *CBS* genes were incomplete, the R_i analyses were interpretable for 165 of the 171 missense changes adjacent to these sites.

Splicing mutations at the presumed branchpoint sequence reduced (mutation 32) or completely abolished (mutation 25) the use of the corresponding acceptor site but did not alter the information content. This was not surprising because this signal is not detectable by information analysis[3] or by other methods. However, one of these branchpoint mutations was also predicted to activate cryptic splicing (mutation 32).

Relationship of R_i to Phenotype

To determine whether disease phenotype was related to predicted splice-site use, the degrees of dyslipidemia, protein expression, and clinical severity were graded and compared with the R_i values of mutant splice sites. Severe (mutations 1, 13, 16, 21, and 22) and average (mutations 2 to 4, 9 to 12, 14, 18, and 19) dyslipidemia were present in individuals with primary splice-site mutations with R_i values <2.4 bits, but dyslipidemia was not present in individuals with mutant sites ≥2.4 bits. Conversely, mutant sites with R_i values ≥2.4 bits were found in patients with mild (mutations 8 and 20) or average (mutations 7, 23, and 24) dyslipidemia but not in those with severe dyslipidemia.

Average (mutations 2, 9, 11, 16, 19, 21, and 22), severe (mutation 1), and mild (mutation 18, which is borderline average) decreases in protein levels or enzymatic activity were found in individuals harboring mutant splice sites <2.4 bits (Table 3). Mutations that resulted in sites with R_i≥2.4 bits, however, exhibited only mild (mutations 8 and 20) or average (mutation 7) reductions in protein expression.

On the basis of age at onset of coronary symptoms, the severity of clinical atherosclerosis for one patient who had R_i<2.4 bits was graded as average (mutation 3); it was graded mild for another individual carrying a mutation with R_i>2.4 bits (mutation 20). Corresponding clinical data were not available for the other mutations.

Discussion

The decrease in information content at mutant splice sites in atherosclerosis candidate genes is related to diminished splice-site use. Predicted reductions in the proportions of normal mRNA were confirmed for 11 splice-site mutations. The decrease in R_i also tended to be related to the degree of reduced protein expression and to phenotypic severity (Table 3).

Missense mutations can, in some cases, simultaneously affect both translation and splicing, and the splicing effects can be detected by information analysis.[4,7] We found 2 missense changes within codons adjacent to splice junctions (mutations 23 and 24) that were predicted to impair splicing. By contrast, 5 other similarly located missense mutations slightly increased (mutations 27 to 29) or decreased (mutations 30 and 31) R_i values without affecting splicing.

Incomplete genomic sequences for splice sites in the *LDLR* and *CBS* genes were analyzed by comparing the ranges of possible R_i values for the corresponding natural and variant sites. The results were interpretable for all except 6 mutations in which the R_i intervals of natural and variant acceptor sites overlapped. Thus, the 5′ and 3′ terminal intronic sequences (positions −16 to −25 of acceptor and +6 of donor sites) can sometimes make a significant contribution to the overall information content of a

splice site.

The decrease in R_i values for splicing mutations was related to the severity of dyslipidemia (Table 3). Splicing mutations that resulted in milder dyslipidemia were predicted to be leaky ($\geq$2.4 bits), whereas severely affected individuals carried mutations that presumably inactivated splice sites and abolished protein expression ($<$2.4 bits). The phenotypes at the outlying lipid values were concordant with these predictions, although corresponding R_i values for mutations in individuals with average dyslipidemia were not predictive. More refined models, incorporating lipid phenotypes as quantitative traits, will require analysis of additional patients and mutations.

Acknowledgments

We thank Drs Thomas D. Schneider and Christoph A. Nienaber for their suggestions. Support was provided by grants from the Public Health Service (CA74683), the American Cancer Society (DHP-132), Merck Genome Research Foundation (to P.K.R.), and Deutsche Forschungsgemeinschaft (KO 1828/1 and KO 1828/1-2; to Y.v.K).

References

1. Mehrabian M, Lusis AJ. Genetic markers for studies of atherosclerosis and related risk factors. In: Lusis AJ, Rotter JI, Sparkes RS, eds. *Molecular Genetics of Coronary Artery Disease*: *Candidate Genes and Process in Atherosclerosis*. Vol 14. Basel: Karger; 1992;363–418.
2. Rogan PK, Faux BM, Schneider TD. Information analysis of human splice site mutations. *Hum Mutat*. 1998;12:153–171.
3. Stephens RM, Schneider TD. Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. *J Mol Biol*. 1992;228:1124–1136.
4. Schneider TD. Information content of individual genetic sequences. *J Theor Biol*. 1997;189:427–441.
5. Krawczak M, Cooper DN. The human gene mutation database. *Trends Genet*. 1997;13:121–122.
6. Reed R, Maniatis T. A role for exon sequences and splice site proximity in splice site selection. *Cell*. 1996;46:681–690.
7. Schneider TD. Sequence walkers: a graphical method to display how binding proteins interact with DNA or RNA sequences. *Nucleic Acids Res*. 1997;25:4408–4415.
8. Rogan PK, Schneider TD. Using information content and base frequencies to distinguish mutations from genetic polymorphisms in splice junction recognition sites. *Hum Mutat*. 1995;6:74–76.
9. Sakai N, Santamarina-Fojo S, Yamashita S, Matsuzawa Y, Brewer HB Jr. Exon 10 skipping caused by intron 10 splice donor site mutation in cholesteryl ester transfer protein gene results in abnormal downstream splice site selection. *J Lipid Res*. 1996;37:2065–2073.
10. Chimienti G, Capurso A, Resta F, Pepe G. G→C change at the donor splice site of intron 1 causes lipoprotein lipase deficiency in a southern-Italian family. *Biochem Biophys Res Commun*. 1992;187:620–627.
11. Sun XM, Patel DD, Bhatnagar D, Knight BL, Soutar AK. Characterization of a splice-site mutation in the gene for the LDL receptor associated with an unpredictably severe clinical phenotype in English patients with heterozygous FH. *Arterioscler Thromb Vasc Biol*. 1995;15:219–227.
12. Jensen HK, Jensen LG, Hansen PS, Bolund L, Færgeman O, Gregersen N. Two point mutations (313+1G→A and 313+1G→T) in the splice donor site of intron 3 of the low-density lipoprotein receptor gene are associated with familial hypercholesterolemia. *Hum Mutat*. 1996;7:269–271.
13. Gudnason V, Sigurdsson G, Nissen H, Humphries SE. Common founder mutation in the LDL receptor gene causing familial hypercholesterolemia in the Iceland population. *Hum Mutat*. 1997;10:36–44.
14. Pagani F, Garcia R, Pariyarath R, Stuani C, Gridelli B, Paone G, Baralle FE. Expression of lysosomal acid lipase mutants detected in three patients with cholesteryl ester storage disease. *Hum Mol Genet*. 1996;5:1611–1617.
15. Muntoni S, Wiebusch H, Funke H, Ros E, Seedorf U, Assmann G. Homozygosity for a splice junction mutation in exon 8 of the gene encoding lysosomal acid lipase in a Spanish kindred with cholesterol ester storage disease (CESD). *Hum Genet*. 1995;95:491–494.
16. Klima H, Ullrich K, Aslanidis C, Fehringer P, Lackner KJ, Schmitz G. A splice junction mutation causes deletion of a 72-base exon from the mRNA for lysosomal acid lipase in a patient with cholesteryl ester storage disease. *J Clin Invest*. 1993;92:2713–2718.
17. Hölzl B, Huber R, Paulweber B, Patsch JR, Sandhofer F. Lipoprotein lipase deficiency due to a 3′ splice site mutation in intron 6 of the lipoprotein lipase gene. *J Lipid Res*. 1994;35:2161–2169.
18. Deeb SS, Takata K, Peng R, Kajiyma G, Albers JJ. A splice-junction mutation responsible for familial apolipoprotein A-II deficiency. *Am J Hum Genet*. 1990;46:822–827.
19. Huang L-S, Kayden H, Sokol RJ, Breslow JL. *APOB* gene nonsense and splicing mutations in a compound heterozygote for familial hypobetalipoproteinemia. *J Lipid Res*. 1991;32:1341–1348.
20. Fojo SS, Beisiegel U, Beil U, Higuchi K, Bojanovski M, Gregg RE, Greten H, Brewer HB. Donor splice site mutation in the apolipoprotein (APO) C-II gene ($APOCII_{Hamburg}$) in a patient with *APOCII* deficiency. *J Clin Invest*. 1988;82:1489–1494.
21. Hata A, Emi M, Luc G, Basdevant A, Gambert P, Iverius P-H, Lalouel J-M. Compound heterozygote for lipoprotein lipase deficiency: Ser->Thr^{244} and transition in 3′ splice site of intron 2 (AG→AA) in the lipoprotein lipase gene. *Am J Hum Genet*. 1990;47:721–726.
22. Top B, van der Zee A, Havekes LM, van't Hooft, Frants RR. Identification of a splice-site mutation in the low density lipoprotein receptor gene by denaturing gradient gel electrophoresis. *Hum Genet*. 1993;91:480–484.
23. Ekström U, Abrahamson M, Sveger T, Lombardi P, Nilsson-Ehle P. An efficient screening procedure detecting six novel mutations in the LDL receptor gene in Swedish children with hypercholesterolemia. *Hum Genet*. 1995;96:147–150.
24. Leren TP, Solberg K, Rødningen OK, Tonstad S, Ose L. Two novel point mutations in the EGF precursor homology domain of the LDL receptor gene causing familial hypercholesterolemia. *Hum Genet*. 1995;96:241–242.
25. Maruyama T, Miyake Y, Tajima S, Harada-Shiba M, Yamamura T, Tsushima M, Kishino B, Horiguchi Y, Funahashi T, Matsuzawa Y. Common mutations in the low-density-lipoprotein-receptor gene causing familial hypercholesterolemia in the Japanese population. *Arterioscler Thromb Vasc Biol*. 1995;15:1713–1718.
26. Funahashi T, Yamashita S, Maruyama T, Ueyama Y, Menju M, Nagali Y, Takemura K, Miyake Y, Tajima S, Matsuzawa Y. A compound heterozygote for familial hypercholesterolemia with homozygous mother. *J Intern Med*. 1996;239:187–190.
27. Gotoda T, Senda M, Murase T, Yamada N, Takaku F, Furuichi Y. Gene polymorphism identified by PvuII in familial lipoprotein lipase deficiency. *Biochem Biophys Res Commun*. 1989;164:1391–1396.
28. Gotoda T, Yamada N, Kawamura M, Kozaki K, Mori N, Ishibashi S, Shimano H, Takaku F, Yazaki Y, Furuichi Y, Murase T. Heterogeneous mutations in the human lipoprotein lipase gene in patients with familial lipoprotein lipase deficiency. *J Clin Invest*. 1991;88:1856–1864.
29. Lelli N, Garuti R, Ghisellini M, Tiozzo R, Rolleri M, Aimale V, Ginocchio E, Naselli A, Bertolini S, Calandra S. Occurrence of multiple aberrantly spliced mRNAs of the LDL-receptor gene upon a donor splice site mutation that causes familial hypercholesterolemia (FHBenevento). *J Lipid Res*. 1995;36:1315–1324.
30. Cladaras C, Hadzopoulou-Cladaras M, Felber BK, Pavlakis G, Zannis VI. The molecular basis of a familial *APOE* deficiency: an acceptor splice site mutation in the third intron of the deficient *APOE* gene. *J Biol Chem*. 1987;262:2310–2315.
31. Schaefer EJ, Gregg RE, Ghiselli G, Forte TM, Ordovas JM, Zech LA, Bewer HB Jr. Familial apolipoprotein E deficiency. *J Clin Invest*. 1986;78:1206–1219.
32. Talmud PJ, Krul ES, Pessah M, Gay G, Schonfeld G, Humphries SE, Infante R. Donor splice mutation generates a lipid-associated apolipoprotein B-27.6 in a patient with homozygous hypobetalipoproteinemia. *J Lipid Res*. 1994;35:468–477.
33. Brown ML, Inazu A, Hesler CB, Agellon LB, Mann C, Whitlock ME, Marcel YL, Milne RW, Koizumi J, Mabuchi H, Takeda R, Tall AR. Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature*. 1989;342:448–451.
34. Koizumi J, Mabuchi H, Yoshimura A, Michishita I, Takeda M, Itoh H, Sakai Y, Ueda K, Takeda R. Deficiency of serum cholesteryl-ester transfer activity in patients with familial hyperalphalipoproteinemia. *Atherosclerosis*. 1985;58:175–186.

35. Koivisto U-M, Viikari JS, Kontula K. Molecular characterization of minor gene rearrangements in Finnish patients with heterozygous familial hypercholesterolemia: identification of two common missense mutations (Gly823→Asp and Leu380→His) and eight rare mutations of the LDL receptor gene. *Am J Hum Genet*. 1995;57:789–797.
36. Pereira E, Ferreira R, Hermelin B, Thomas G, Bernard C, Bertrand V, Nassiff H, Del Castillo DM, Bereziat G, Benlian P. Recurrent and novel LDL receptor gene mutations causing heterozygous familial hypercholesterolemia in La Habana. *Hum Genet*. 1995;96:319–322.
37. Kuivenhoven JA, Weibusch H, Pritchard PH, Funke H, Benne R, Assmann G, Kastelein JJP. An intronic mutation in a lariat branchpoint is a direct cause of an inherited human disorder (fish eye disease). *J Clin Invest*. 1996;98:358–364.
38. Jensen HK, Jensen LG, Hansen PS, Faergeman O, Gregersen N. High sensitivity of the single-strand conformation polymorphism method for detecting sequence variations in the low-density lipoprotein receptor gene validated by DNA sequencing. *Clin Chem*. 1996;42:1140–1146.
39. Sundvold H, Solberg K, Tonstad S, Rødningen OK, Ose L, Berg K, Leren TP. A common missense mutation (C210G) in the LDL receptor gene among Norwegian familial hypercholesterolemia subjects. *Hum Mutat*. 1996;7:70–71.
40. Day INM, Whittall RA, O'Dell SD, Haddad L, Bolla MK, Gudnason V, Humphries SE. Spectrum of LDL receptor gene mutations in heterozygous familial hypercholesterolemia. *Hum Mutat*. 1997;10:116–127.
41. Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat*. 1992;1:445–466.
42. Bruin T, Tutzgöl S, van Diermen DE, Hoogerbrugge-van der Linden N, Brunzell JD, Hayden MR, Kastelein JJP. Recurrent pancreatitis and chylomicronemia in an extended Dutch kindred is caused by Gly15→Ser substitution in lipoprotein lipase. *J Lipid Res*. 1993;34:2109–2119.
43. Webb JC, Patel DD, Shoulders CC, Knight BL, Soutar AK. Genetic variation at a splicing branch point in intron 9 of the low density lipoprotein(LDL)-receptor gene: a rare mutation that disrupts mRNA splicing in a patient with familial hypercholesterolemia and a common polymorphism. *Hum Mol Genet*. 1996;5:1325–1331.
44. Heidberg A. The risk of atherosclerotic vascular disease in subjects with xanthomatosis. *Acta Med Scand*. 1975;198:249–261.
45. Morganroth J, Levy RI, Fredrickson DS. The biochemical, clinical, and genetic features of type III hyperlipoproteinemia. *Ann Intern Med*. 1975;82:158–174.