J Hum Genet (2001) 46:522–528

ORIGINAL ARTICLE

Aritoshi Iida · Susumu Saito · Akihiro Sekine
Yuri Kitamura · Kimie Kondo · Chihiro Mishima
Saori Osawa · Satoko Harigae · Yusuke Nakamura

High-density single-nucleotide polymorphism (SNP) map of the 150-kb region corresponding to the human ATP-binding cassette transporter A1 (*ABCA1*) gene

Received: May 22, 2001 / Accepted: June 12, 2001

Abstract Highly dense catalogs of human genetic variations, in combination with high-throughput genotyping technologies, are expected to clarify individual genetic differences in pharmacological responsiveness and predispositions to common diseases. Here we report single-nucleotide polymorphisms (SNPs) present among 48 Japanese individuals at the locus for the human ATP-binding cassette transporter A1 (*ABCA1*) gene. ABCA1 plays a key role in apolipoprotein-mediated cholesterol transport, and mutations in this gene are responsible for Tangier disease and familial high-density lipoprotein deficiency associated with reduced cholesterol efflux. We identified a total of 162 SNPs, 149 of which were novel, within the 150-kb region encompassing the entire *ABCA1* gene. Eight of the SNPs lie within coding elements, two in 5′ flanking regions, 147 in introns, and five in 3′ untranslated regions, but none were found in 5′ untranslated or 3′ flanking regions. The ratio of transitions to transversions was approximately 2.37 to 1. Our dense SNP map of this region could serve as a powerful resource for studies of complex genetic diseases that may be associated with *ABCA1* and of individual responses to drug therapy.

Key words Single-nucleotide polymorphism (SNP) · Insertion-deletion polymorphism · High-density SNP map · ATP-binding cassette A1 transporter gene · Japanese population · Nonsynonymous substitution

A. Iida · S. Saito · A. Sekine · Y. Kitamura · K. Kondo · C. Mishima · S. Osawa · S. Harigae · Y. Nakamura
Laboratory for Genotyping, RIKEN SNP Research Center, Tokyo, Japan

Y. Nakamura (✉)
Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Tel. +81-3-5449-5372; Fax +81-3-5449-5433
e-mail: yusuke@ims.u-tokyo.ac.jp

Introduction

The human ATP-binding cassette transporter A1 (*ABCA1*) gene encodes a 247-kDa membrane protein with two separate transmembrane portions, each consisting of six membrane-spanning domains and a nucleotide-binding fold. The regulatory domain contains a highly hydrophobic segment and is expressed on the plasma membrane and the Golgi complex (Luciani et al. 1994; Orso et al. 2000). Functional analyses of ABCA1 have revealed roles in (1) regulation of the apolipoprotein A1-dependent cellular export of cholesterol and phopholipids (Langmann et al. 1999; Oram 2000); (2) engulfment of apoptotic cells by macrophages (Luciani and Chimini 1996; Hamon et al. 2000); (3) secretion of macrophage interleukin-1 beta (Hamon et al. 1997); and (4) caveolar processing (Orso et al. 2000). The genomic sequence at the *ABCA1* locus that has been determined so far includes 1453 bp of the promoter region, 146,581 bp of introns and exons, and 1000 bp of DNA downstream of the polyadenylation signal (Santamarina-Fojo et al. 2000). The gene itself consists of 50 exons ranging from 33 to 245 bp. Relationships between *ABCA1* gene and human dyslipidemic diseases have been investigated intensively; mutations of *ABCA1* gene can be found in patients with Tangier disease and in persons with familial high-density lipoprotein (HDL) deficiency associated with reduced cholesterol efflux (Bodzioch et al. 1999; Brooks-Wilson et al. 1999; Marcil et al. 1999; Rust et al. 1999). Recently Clee et al. (2001) isolated 16 single-nucleotide polymorphisms (SNPs) from coding elements of the *ABCA1* gene and revealed that some are associated with altered plasma lipid levels and risk of coronary artery disease.

Variations in genes that encode drug-transporter and drug-metabolizing enzymes may be associated with susceptibility to common diseases and with differences in therapeutic efficacy and side effects of drugs among individual persons, because some variations can alter the activity of a gene product with respect to quality, quantity, or both. Therefore, information concerning naturally occurring genetic variants in human transporter genes such as *ABCA1*

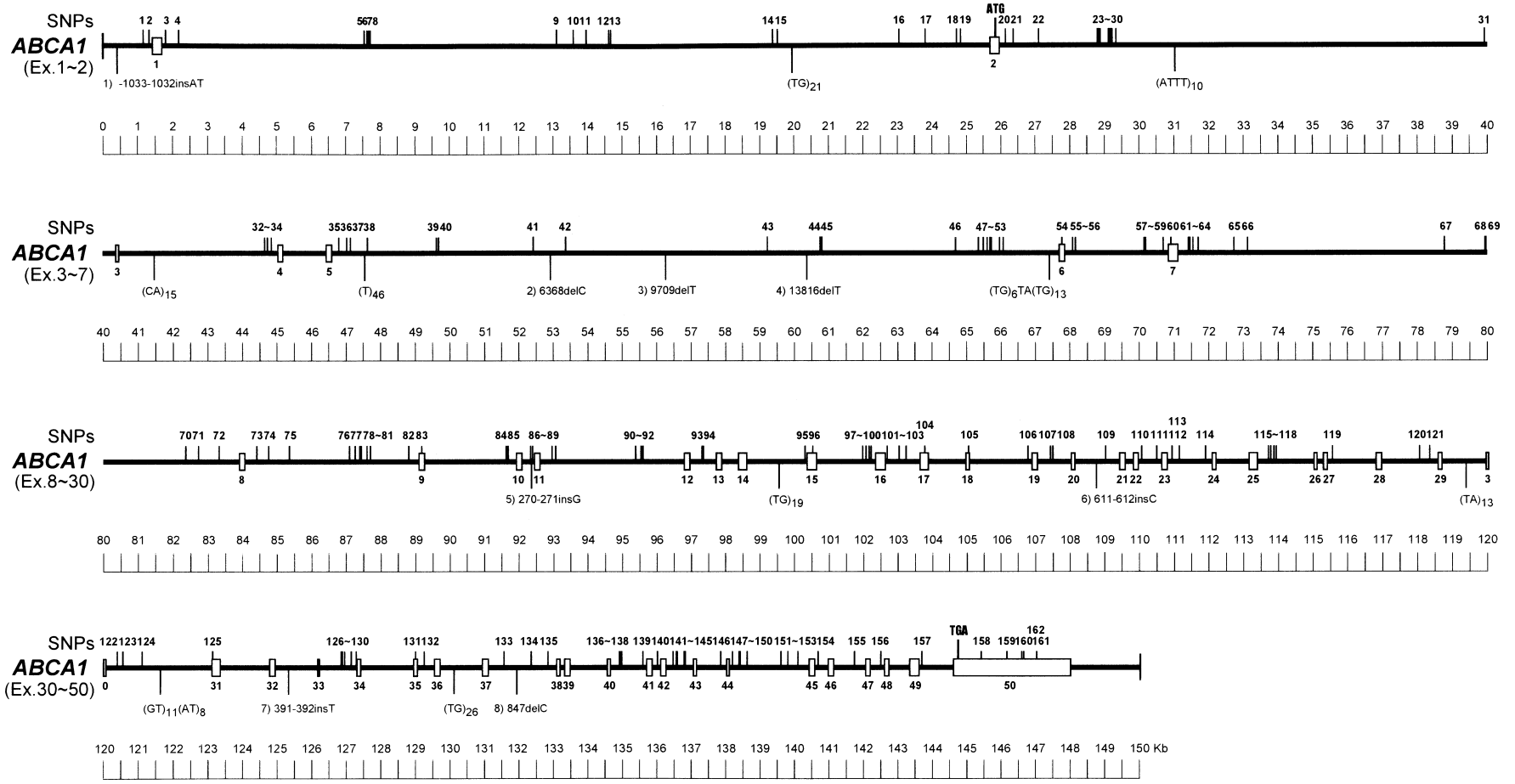

Fig. 1. Single-nucleotide polymorphism (SNP) map spanning the 150-kb region containing the *ABCA1* gene. Exons are represented by *open rectangles* and introns by *horizontal lines*. SNPs are indicated above the lines according to number (corresponding to the numbers in the far-left column of Table 1); the positions of eight insertion-deletion polymorphisms are indicated *below the lines* (see Table 2). Microsatellite sequences are also shown

Table 1. Characterization of 162 single-nucleotide polymorphisms (SNPs) within the *ABCA1* locus

Number	Location	Exon	SNP (5′ to 3′)	Substituion	Repetitive sequence	Identity to dbSNP	Reference
1	−278 G > C 5′ flanking region		gggcccgggcgggggaaggg G/C acgcagaccgcggaccctaa			rs1800976	
2	−99 G > C 5′ flanking region		acataaacagaggccgggaa G/C ggggcggggaggagggagag				
3	159 G > T intron 1		gcggtgttaaatggggagac G/T atgtcctagtacgagctctg				
4	506 G > C intron 1		gaattggctatatgctcccc G/C ggactggagcggcacagtcc				
5	5897 T > G intron 1		gtacaaaaacctttagcttt T/G gcaaacctcctttaagaccc				
6	5929 C > T intron 1		ttaagacccgatttaaatgc C/T tccctcctcatgaagctctt				
7	5962 T > C intron 1		aagctcttctggatccactc T/C ttcccatcactaagttgaaa				
8	5985 A > C intron 1		cccatcactaagttgaaagt A/C agatccccttctctttactt				
9	11416 G > A intron 1		ttacagtgccctttatagga G/A agaaagaagaaattgtgtct				
10	11935 G > A intron 1		tctctgtggagcaaatagag G/A gctgtctgacacttggttcc				
11	12281 T > A intron 1		gaatgtttgatttgtgaaaa T/A cttaataacagtagttttt				
12	12924 T > C intron 1		gtgctgacaatcttatactc T/C aggttgaacctccggggaag				
13	13002 C > G intron 1		gagcctcaatcacagattct C/G tctagctcacatgaagttaa				
14	17715 C > T intron 1		ggagcatgactttgtggaag C/T ctctcctcttccacccagag				
15	17848 T > C intron 1		gagggctgactgtcaccctt T/C gataggagcccagcactaaa				
16	21384 G > C intron 1		gtgggtgggaggaattggag G/C aggaagcttgcctaagtgtg				
17	22145 C > G intron 1		gtagcttctaaatcaacgaa C/G tgattcctggagagcagctt			rs1340361	
18	23063 G > A intron 1		ggaggcacctgtgacaccca G/A cggagtaggggggcggtgtg				
19	23131 G > A intron 1		agtgtgcatatgtgctgacc G/A tgggagcttgtttgtcggtt				
20	156 T > C intron 2		ggacacaggactgtgtggtc T/C ggatatggcatgtggcttat			rs1078143	
21	384 A > G intron 2		gctgtgggtgaagtgagtta A/G tggccccactcttagagatc			rs1078144	
22	1081 G > A intron 2		agtgcagccaaaattgcaaa G/A tcataccattcaaattaata			rs752187	
23	2801 A > G intron 2		aagaaaagtgatttatttca A/G gttgctgatgcttagattgt				
24	2830 C > G intron 2		tgcttagattgttagagttg C/G aaagatctggcttgcatctt				
25	2856 A > G intron 2		tctggcttgcatcttgtaca A/G ctgacagaactggggctcag				
26	3187 A > G intron 2		tgatagctgttgcctgcagc A/G tacggacgttcattgcgcag				
27	3190 C > T intron 2		tagctgttgcctgcagcata C/T ggacgttcattgcgcagttc				
28	3194 C > T intron 2		tgttgcctgcagcatacgga C/T gttcattgcgcagttcctgt				
29	3204 G > A intron 2		agcatacggacgttcattgc G/A cagttcctgtctcctgagat				
30	3401 T > C intron 2		acataaagcctgtgtgctgc T/C gccaggaagactagaaacgc				
31	13927 A > G intron 2		gtcaccacatacctggcact A/G tgctaaggctgggaatgcag		L2		
32	4163 G > A intron 3		ccagcccacttcatcttacc G/A tagttacctccttagagtat				
33	4262 T > C intron 3		tgtcaaagaggaactaagga T/C gccagggactttctgcttag				
34	4306 C > T intron 3		ccctctcatcacttctccaa C/T gctggtatcatgaaccccat				
35	240 G > A intron 5		gacagaagaaaagtccccag G/A gaagaatactacagacttgg			rs1107281	
36	490 G > A intron 5		gatgggcatttgaacttgtt G/A tctttaaaaagtgaaatctt				
37	583 T > G intron 5		tatctgggagtgggcattt T/G ctgactgaggcattggctgc				
38	1051 C > T inton 5		ggctacaaaactgtgctttc C/T ttgggcagtaaaagaggcaa				
39	3051 G > A intron 5		tagagaacaagtctaattct G/A ttttccttgaaatagtcgaa				
40	3127 A > G intron 5		aagtccatgatttttaggc A/G aaatggcctcctttcctctt				
41	5924 C > T intron 5		ctttctttcacaaaaattgcc C/T cccagagctttctggaaggg				
42	6831 T > C intron 5		ccagtccctcagccttgcca T/C tgcttatgctggtctggaaa				
43	12678 G > C intron 5		gctcaccgctctgctcaccc G/C accctctggccatctcctct				
44	14214 G > A intron 5		cagcttggtcccagaggcct G/A gacctgggtcccagaggtcc				
45	14257 C > T intron 5		gctggttccccggcttggtc C/T cagaggcctggatgtgtggc				
46	18078 C > T intron 5		cctaccacaccatgcacgtg C/T acagccaagggttgttgact				
47	18795 G > A intron 5		ctgggctcttcctggacctg G/A ccagctaaaaggaaatctcc				
48	18948 G > A intron 5		gcattggtggtactaagaac G/A catattccctatcctatagg		L2		
49	19053 T > C intron 5		ctcccccaacattaaaagtg T/C aagggatgcttattcaaatg		MER5A		
50	19148 C > A intron 5		ggcccaagaaactgcatttt C/A gcatgctccctaaatgaagc		MER5A		
51	19229 C > T intron 5		atgctaacagtgtagagtca C/T atgtgatgggaagcatcagg				
52	19405 T > C intron 5		cttgctcaatttattctgtc T/C atataactcaatattactga				
53	19534 G > A intron 5		catgtgaccctcttagctcc G/A cggattaactcctgtcctca				
54	474 G > A coding region	6	gaaaccttctctgggttcct G/A tatcacaacctctctctccc	Leu 158 Leu			
55	210 A > C intron 6		gcaacctggcgtcatgggcc A/C gctggttaaaataaaattga				
56	334 G > A intron 6		acagttctgaggcaataacc G/A tggttaagggttattgatct				
57	2288 C > T intron 6		cttctttcaaagcttgtggt C/T cactggaccacgtatgaagt				
58	2322 T > C intron 6		atgaagtagaatagtttagg T/C ccagaaaggcaattaagtaa				
59	2820 T > G intron 6		gtgctttgatacattctgag T/G ttcagtaaagagacctgatg				
60	656 G > A coding region	7	tgagctttgtggcctaccaa G/A ggagaaactggctgcagcag	Arg 219 Lys			Clee et al. 2001
61	416 G > A intron 7		catcataaagatgacattgt G/A ggctgtcacagttggaaggc				
62	471 C > T intron 7		agaccacactatttagctta C/T ttagtaataacattgcaaag				
63	504 G > A intron 7		ttgcaaagaaaaattccgac G/A aagtttttcagcctaggaa				
64	679 G > C intron 7		gctctggtgaaattcctctc G/C ctaccccaaacatcatcatt				
65	1740 C > T intron 7		acaaatgctcaccctttcag C/T tggaatgattgaaattttgg				
66	2122 A > G intron 7		tgattaaggtggctactacc A/G ggtgctttctgcatatctcg				
67	7753 T > C intrion 7		taggaattccaagctgtgaa T/C tttttactgaagctctttgg				
68	8973 A > T intron 7		atggaaatttgtttatattg A/T ctacagattgccaatattat				
69	8976 A > G intron 7		gaaatttgtttatattgact A/G cagattgccaatattattag				

Table 1. Continued

Number	Location	Exon	SNP (5′ to 3′)	Substituion	Repetitive sequence	Identity to dbSNP	Reference
70	11327 G > C intron 7		ctaacaatcttatttccatt G/C agtccttataaaagaagtgg				
71	11738 C > T intron 7		ctgacgtttaagggagaccg C/T gtaggtccctttgaggactg				
72	12295 T > A intron 7		agtctgtaaattattgttct T/A ttttttctttagcttatgct				
73	387 C > G intron 8		tagcaaggccaatcatttta C/G caacacacatgcttgctaac				
74	697 A > T intron 8		ggaactgtctggtgtccccc A/T gcataggaagctgagccagg				
75	1312 G > A intron 8		attgctctgcagatcccctc G/A cagccctctgtcccttgttc			rs1175929	
76	3036 T > G intron 8		ctttatgtgggaagaaattt T/G ttttttgattggggagtgg				
77	3176 C > A intron 8		aaatggcctggttctctgtc C/A cctttctgtctgtatgcctc				
78	3364 A > T intron 8		ggcagaaggcaaagcttagg A/T cctagagagtgctggaccac				
79	3373 G > A intron 8		caaagcttaggacctagaga G/A tgctggaccacgccactcac				
80	3561 C > A intron 8		cagggatttattaatgattt C/A ttgtgaaatgtttggaaata				
81	3654 T > C intron 8		agtgccggaatacatttgca T/C gtaagacagaacgctgcctg				
82	4715 C > T intron 8		ggcagagggctctcagaatc C/T gcatttccaacaatgtctcc				
83	936 C > T coding region	9	cgtattgtctgcgggcatcc C/T gagggaggggggctgaagat	Pro 312 Pro			
84	2309 A > G intron 9		cccctcaagagtcagtttaa A/G tgttggtcatgttagttgtc				
85	2392 T > C intron 9		atgggagggcttgtgcttca T/C gaaaacatttttccagatca				
86	228 A > G intron 10		tggggatggggaggactggc A/G cagggctgctgtgatggggt				
87	319 C > T intron 10		ttctgcggtccctggctccc C/T acctgactccaggtgaacaa				
88	377 A > C intron 11		gaaagaagtgtgggagcaaa A/C gcatgatgttacatgtagac				
89	521 G > A intron 11		agtgctctagagacaattgg G/A ttcaaatgtggagcaggctg				
90	2850 G > C intron 11		ctctatacaatcattatgct G/C ccattgaaataataaataca				
91	2976 A > G intron 11		ctccaattcggtagaaccag A/G gcttcatcttctctgtcgaa				
92	3056 C > T intron 11		gtttgcagctgctgtttttc C/T ggcagcacatctgtgcaggc				
93	340 T > C intron 12		ggcattatttgtgaaactta T/C ctaaaatcgaattcgggtcc				
94	381 A > G intron 12		aattaaatttttgaaatttt A/G tattaaaaattatattagta				
95	1728 C > T intron 14		caggctcagaggccttggcc C/T atcaccctggctcacgtgtg				
96	2040 C > A coding region	15	atgggcctggacaacagcat C/A ctctggtttagctggttcat	Ile 680 Ile			
97	1382 G > A intron 15		cttttagacagaaaagttac G/A tgggatattatctcccacag				
98	1453 G > A intron 15		tatataaggagaaaccagtt G/A aaattacctattgaagaaac				
99	1567 G > A intron 15		ttctgcgtagttttgggtaa G/A tcacttatcttctttaggat		MIR		
100	1617 T > A intron 15		cagttgcctcatcagaaaga T/A gaacagcattacgcctctgc		MIR		
101	95 T > A intron 16		agttgagaacagaagatgat T/A gtcttttccaatgggacatg				
102	452 G > A intron 16		tggtgttttgcttgagtaat G/A ttttctgaactaagcacaac				
103	657 T > C intron 16		ctgttgcctcagtctgggct T/C cataggcatcagcagcccca				
104	2473 G > A coding region	17	gcttcaatctcaccacttcg G/A tctccatgatgctgtttgac	Val 825 Ile			Clee et al. 2001
105	2649 A > G coding region	18	ggttccaaccagaagagaat A/G tcagaaagtaagtgctgttg	Ile 883 Met			Clee et al. 2001
106	1730 C > G intron 18		tgaaagttcaagcgcagtgc C/G ctgtgtccttacactccact				
107	426 A > G intron 19		aggaccttacagtgggtagt A/G tcaggaggggtcagggggctg				
108	468 A > G intron 19		aaagcaccagcgttagcctc A/G gtggcttccagcacgattcc				
109	876 C > T intron 20		ccctcctcatctaaagtgaa C/T acatggggctcatgtgcagg				
110	118 T > G intron 22		catgggatactcttctgtta T/G cacagaagagataaagggca				
111	560 G > A intron 22		aaagctttgccattctaggg G/A tcatagccatacagggtgaa				
112	102 A > G intron 23		acccctttgccatgttgaa A/G ccaccatctccctgctctgt				
113	287 C > T intron 23		gtcaaagaaaagagacttgt C/T aagaggtaagagccttggct				
114	1063 G > A intron 23		acctttcaccctcaggaagc G/A aggctgttcacacggcacac				
115	321 T > G intron 25		ctctttacttaagtacagtg T/G gaggaacagcggcatcagga		MER5A		
116	376 G > C intron 25		gttagaaattcagcaacttg G/C gcccagctcagacctactga		MER5A		
117	478 C > T intron 25		catacataggaaatgacaaa C/T gtttatggatggatagtcta				
118	579 G > T intron 25		tcatttaattctcaaaaaaa G/T atgaaaaaatgaacactcag				
119	153 C > T intron 27		aatggtaaaagccacttgtt C/T tttgcagcatcgtgcatgtg				
120	1058 C > T intron 28		actatcatgggagataatga C/T tatggttgtccatgattgga				
121	1317 C > T intron 28		caggacccagtgttctgagt C/T accctgaatgtgagcactat				
122	372 T > C intron 30		tatatgatttttaggttttg T/C ttatcagcttcttcgctttt				
123	506 A > G intron 30		ccttttaaaaagtaagcagt A/G gataaataaattcagtgaag				
124	1033 G > C intron 30		ctggatttcatggtgccttt G/C attttccacatgaaggttgt				
125	4281 G > A coding region	31	tcttccctttgcagagacac G/A ccctgccaggcaggggagga	Thr 1427 Thr			
126	626 C > T intron 33		ggctccttgttactgatttc C/T gtcttttctctctgcctttt				
127	719 G > A intron 33		taatagccctcatgctagaa G/A ggagccggagcctgtgtata				
128	726 G > A intron 33		cctcatgctagaagggagcc G/A gagcctgtgtataaggccag				
129	889 A > G intron 33		ctttcctcaatgtctcagct A/G tctaactgtgtgtgtaatca				
130	1097 G > C intron 33		ctgtgcaccccactgtctgg G/C ttttaatgtcaggctgttct				
131	4760 G > A coding region	35	tatgacaggactggacacca G/A aaataatgtcaaggtaaacc	Arg 1587 Lys			Clee et al. 2001
132	234 T > C intron 35		aacctatctaaacctcagtt T/C cctcatctgtgaaatggaga		MIR		
133	411 C > T intron 37		aactctgtacattttatcag C/T agcttatccatccattgcaa				
134	1224 A > G intron 37		caggcataggtgattcagag A/G tgaaaggtcaagtccctgaa		L2		
135	1720 G > T intron 37		aaattaaaattactctgact G/T ggaatccatcgttcagtaag				
136	251 T > G intron 40		tgaaggtaaggaaaatagtg T/G tatttgcttggatccactgg				
137	252 T > C intron 40		gaaggtaaggaaaatagtgt T/C atttgcttggatccactggc				
138	319 A > G intron 40		agcactggaaaagtcaaacc A/G taactttgagaattaggtga				

Table 1. Continued

Number	Location	Exon	SNP (5′ to 3′)	Substituion	Repetitive sequence	Identity to dbSNP	Reference
139	957 G > C intron 40		cttgttactctttttcctt G/C tcatgggtgatagccatttg				
140	146 C > T intron 41		tgatgtgggcatcccgcagc C/T ccctccctgcccatcctgga				
141	239 A > C intron 42		cattggttttatatgcttac A/C tttatgtgttagttattaaa				
142	321 T > A intron 42		aataaatggttgattttgag T/A ttgagtttcatagtccaaaa				
143	322 T > C intron 42		ataaatggttgattttgagt T/C tgagtttcatagtccaaaaa				
144	533 G > A intron 42		agatgaaaaattatgtagat G/A ataatgaatgatacggttct				
145	546 A > G intron 42		tgtagatgataatgaatgat A/G cggttctaaaaagacaggtt				
146	739 T > A intron 43		tacagccacacttaaaatgg T/A cccattatgaaatacatatt				
147	18 T > C intron 44		taggtgagaaaaagaagtggc T/C tgtattttgctgcaaagact				
148	264 T > C intron 44		acaatataatttgcttgttt T/C ttaagagtataatttagtga		L1MB8		
149	279 T > C intron 44		tgtttttttaagagtataatt T/C agtgatttttggtaaattga		L1MB8		
150	508 C > T intron 44		tttacattgctacataaaat C/T cccctatgtacatgtaccta				
151	1477 A > T intron 44		gatctcctctcctgtctctt A/T cattttgcagtagcaatgt				
152	1665 G > A intron 44		tggttgtaagaactgatttg G/A ttggtatagctgtgagggcc				
153	1956 T > G intron 44		gtgttgctcacactcaaaat T/G tctgggccttctcatttggt				
154	68 T > C intron 45		aatatatacctatggcttt T/C ccacacgcattgacttcagg				
155	608 G > C intron 46		ttatactgacttcaatagag G/C tttcagacaaaaagttgttt				
156	336 T > C intron 47		ttcacaattgtaaacaccac T/C acactgaacagcatcatccc		L1MD2		
157	55 G > C intron 49		agggtgtggattcctgcccc G/C acactcccgcccataggtcc			rs1331924	
158	7479 C > T 3′ untranslated region	50	aacaaaaatgtgggtgtctc C/T aggcacgggaaacttggttc				
159	8226 C > T 3′ untranslated region	50	aggagcccactgtaacaata C/T tgggcagccttttttttttt				
160	8682 G > A 3′ untranslated region	50	aacttcttccacttttttcca G/A aatttgaatattaacgctaa			rs363717	
161	8697 C > T 3′ untranslated region	50	ttccagaatttgaatattaa C/T gctaaaggtgtaagacttca				
162	9097 A > G 3′ untranslated region	50	aactattttgaagaaaacac A/G acattttaatacagattgaa				

Nucleotide numbering is according to the mutation nomenclature (den Dunnen and Antonarakis, 2000)

Table 2. Characterization of insertion/deletion polymorphisms at the *ABCA1* locus

Number	Location	Variations (5′ to 3′)
1	(−1033)–(−1032) ins AT 5′ flanking region	tgacttaaatatttagacat (AT/+) ggtgtgtaggcctgcattcc
2	6368 del C intron 5	ttctgatggggttgttgctg (C/−) tgagaatcatgactgggtgg
3	9709 del T intron 5	cattttctgtctgaaccccc (T/−) cacccattcaggcagctgct
4	13816 del T intron 5	tccctacttctcctttttt (T/−) catttgcctcctccacccac
5	270–271 ins G intron 10	cttttcagggaggagccaaa (G/+) cgctcattgtctgtgcttct
6	611–612 ins C intron 20	tttagcccatcctctccccc (C/+) gccaccctccttattgaggc
7	391–392 ins T intron 32	gagtgccttgggtactctct (T/+) gatgggggactccatgataa
8	847 del C intron 37	gctgtatattgtgaatgtcc (C/−) gttttcaaaagcaaagccaa

Nucleotide numbering is according to the mutation nomenclature (den Dunnen and Antonarakis, 2000)
(+), insertion polymorphism; (−), deletion polymorphism

is an important resource for understanding not only the etiology and risk of some diseases, but also the pharmacokinetics or pharmacodyamics of drugs used to treat them. We have focused on identifying variations in genomic regions that contain drug-metabolizing enzymes, and so far have isolated 1066 SNPs among 71 such loci (lida et al. 2001a,b; Saito et al. 2001; Saito et al., in preparation; Sekine et al. 2001). We report here a total of 162 SNPs spanning the entire *ABCA1* genomic region, detected among 96 chromosomes in a Japanese population sample.

Subjects and methods

DNA samples, PCR, direct sequencing, and detection of SNPs

Blood samples were obtained with informed consent from 48 healthy Japanese individuals for this study, which was approved by the ethical committee of the RIKEN SNP Research Center. On the basis of *ABCA1*-genomic sequences (accession numbers AF275948.1 and AL359846.11), released from the GenBank database, and information published by Remaley et al. (1999) and Santamarina-Fojo et al. (2000), we designed primers to amplify the *ABCA1* gene in its entirety as well as up to 2 kb upstream from the first exon and downstream of the last exon. We eliminated most of the regions corresponding to repetitive sequences predicted by the RepeatMasker program (http://ftp. genome.washington.edu/cgibin/RepeatMasker). Genomic DNA extraction, polymerase chain reaction (PCR) experiments, and DNA sequencing were performed according to methods described previously (Ohnishi et al. 2000). Each PCR was performed using 20 ng of DNA pooled from three individuals. All SNPs detected by the PolyPhred computer program (Nickerson et al. 1997) were confirmed by sequencing both strands of each PCR product.

Results and discussion

We present here a highly dense SNP map covering the 150-kb genomic region that contains the entire human *ABCA1* gene. A total of 91 partially overlapping PCR fragments, covering 57.4% of the 150-kb *ABCA1* gene locus, were sequenced using 364 primers. The organization of the *ABCA1* gene and locations of identified SNPs are illustrated schematically in Fig. 1. A total of 162 SNPs were identified: eight were located in coding elements, two in the 5′ flanking region, 147 in introns, and five in the 3′ untranslated region (Table 1; also see Fig. 1). No SNP was detected in the 5′ untranslated or 3′ flanking regions. Comparison of our data with SNPs deposited in the dbSNP database in the National Center for Biotechnology Information (NCBI, U.S.) and an earlier report (Clee et al. 2001) indicated that 149 (94%) of the SNPs identified in this study could be considered novel. As a whole, 154 of the 162 SNPs identified (95.0%) were located within 79.4 kb of DNA representing noncoding exons, introns, and resequenced flanking regions (1SNP/516 bp), whereas only 8 SNPs were present in 6783 bp of the coding DNA screened (1SNP/848 bp). Therefore, the frequency of SNPs in noncoding portions of the *ABCA1* gene was 1.64-fold higher than in coding elements. Frequencies of the substitutions were 35.8% for A/G, 34.6% for C/T, 11.7% for C/G, 6.8% for G/T, 6.2% for A/T, and 4.9% for A/C. The ratio of transitions to transversions was approximately 2.37 to 1. The higher prevalence of A/G substitutions and the transition/transversion ratio are both in close accord with previous observations (Cambien et al. 1999; Iida et al. 2001a,b; Venter et al. 2001). In addition to SNPs, we identified eight novel insertion/deletion polymorphisms within introns of the *ABCA1* gene (Table 2).

Among 13 SNPs identified in exonic regions, the 4 nonsynonymous sites that we found in coding elements had been reported previously (Clee et al. 2001). Although the frequency of the coding SNPs, particularly nonsynonymous ones, was very small, they would be likely to influence protein function. In fact, it is reported that the R219K substitution is associated with a decreased severity of atherosclerosis, a decreased risk of coronary events, decreased triglycerides, and a higher level of HDL cholesterol (HDL-C) in plasma (Clee et al. 2001). Moreover, since the regulatory elements for gene expression are contained in 5′ promoter regions and introns and occasionally in 3′ flanking regions, SNPs within those regions may affect the quantity of the gene product. In light of recent progress in understanding the biomedical features of *ABCA1*, the SNPs documented here will be useful not only for studying associations between specific SNPs and some coronary diseases, but also for determining the applicability of pharmacogenomic information to medical practice.

References

Bodzioch M, Orso E, Klucken J, Langmann T, Bottcher A, Diederich W, Drobnik W, Barlage S, Buchler C, Porsch-Ozcurumez M, Kaminski WE, Hahmann HW, Oette K, Rothe G, Aslanidis C, Lackner KJ, Schmitz G (1999) The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. Nat Genet 22:347–351

Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K, van Dam M, Yu L, Brewer C, Collins JA, Molhuizen HO, Loubser O, Ouelette BF, Fichter K, Ashbourne-Excoffon KJ, Sensen CW, Scherer S, Mott S, Denis M, Martindale D, Frohlich J, Morgan K, Koop B, Pimstone S, Kastelein JJ, Genest J Jr, Hayden MR (1999) Mutations in *ABC1* in Tangier disease and familial high-density lipoprotein deficiency. Nat Genet 22:336–345

Cambien F, Poirier O, Nicaud V, Herrmann SM, Mallet C, Ricard S, Behague I, Hallet V, Blanc H, Loukaci V, Thillet J, Evans A, Ruidavets JB, Arveiler D, Luc G, Tiret L (1999) Sequence diversity in 36 candidate genes for cardiovascular disorders. Am J Hum Genet 65:183–191

Clee SM, Zwinderman AH, Engert JC, Zwarts KY, Molhuizen HO, Roomp K, Jukema JW, van Wijland M, van Dam M, Hudson TJ, Brooks-Wilson A, Genest J, Jr, Kastelein JJ, Hayden MR (2001) Common genetic variation in *abca1* is associated with altered lipoprotein levels and a modified risk for coronary artery disease. Circulation 103:1198–1205

den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mut 15:7–12

Hamon Y, Luciani MF, Becq F, Verrier B, Rubartelli A, Chimini G (1997) Interleukin-1beta secretion is impaired by inhibitors of the Atp binding cassette transporter, *ABC1*. Blood 90:2911–2915

Hamon Y, Broccardo C, Chambenoit O, Luciani MF, Toti F, Chaslin S, Freyssinet JM, Devaux PF, McNeish J, Marguet D, Chimini G (2000) *ABC1* promotes engulfment of apoptotic cells and transbilayer redistribution of phosphatidylserine. Nat Cell Biol 2:399–406

Iida A, Sekine A, Saito A, Kitamura Y, Kitamoto T, Osawa S, Mishima C, Nakamura Y (2001a) Catalog of 320 single nucleotide polymorphisms (SNPs) in 20 quinone oxidoreductase and sulfotransferase genes. J Hum Genet 46:225–240

Iida A, Saito S, Sekine A, Kitamoto T, Kitamura Y, Mishima C, Osawa S, Kondo K, Harigae S, Nakamura Y (2001b) Catalog of 434 single nucleotide polymophisms (SNPs) in the alcohol dehydrogenase genes, glutathione S transferase genes, and NADH ubiquinone oxidoreductase genes. J Hum Genet 46:385–407

Langmann T, Klucken J, Reil M, Liebisch G, Luciani MF, Chimini G, Kaminski WE, Schmitz G (1999) Molecular cloning of the human ATP-binding cassette transporter 1 (*hABC1*): evidence for sterol-dependent regulation in macrophages. Biochem Biophys Res Commun 1999 257:29–33

Luciani MF, Chimini G (1996) The ATP binding cassette transporter *ABC1*, is required for the engulfment of corpses generated by apoptotic cell death. EMBO J 15:226–235

Luciani MF, Denizot F, Savary S, Mattei MG, Chimini G (1994) Cloning of two novel ABC transporters mapping on human chromosome 9. Genomics 21:150–159

Marcil M, Brooks-Wilson A, Clee SM, Roomp K, Zhang LH, Yu L, Collins JA, van Dam M, Molhuizen HO, Loubster O, Ouellette BF, Sensen CW, Fichter K, Mott S, Denis M, Boucher B, Pimstone S, Genest J, Jr, Kastelein JJ, Hayden MR (1999) Mutations in the *ABC1* gene in familial HDL deficiency with defective cholesterol efflux. Lancet 354:1341–1346

Nickerson DA, Tobe VO, Taylor SL (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res 25:2745–2751

Ohnishi Y, Tanaka T, Yamada R, Suematsu K, Minami M, Fujii K, Hoki N, Kodama K, Nagata S, Hayashi T, Kinoshita N, Sato H, Sato H, Kuzuya T, Takeda H, Hori M, Nakamura Y (2000) Identification of 187 single nucleotide polymorphisms (SNPs) among 41 candidate genes for ischemic heart disease in the Japanese population. Hum Genet 106:288–292

Oram JF (2000) Tangier disease and ABCA1. Biochim Biophys Acta 1529:321–330

Orso E, Broccardo C, Kaminski WE, Bottcher A, Liebisch G, Drobnik W, Gotz A, Chambenoit O, Diederich W, Langmann T, Spruss T, Luciani MF, Rothe G, Lackner KJ, Chimini G, Schmitz G (2000) Transport of lipids from golgi to plasma membrane is defective in tangier disease patients and *Abc1*-deficient mice. Nat Genet 24:192–196

Remaley AT, Rust S, Rosier M, Knapper C, Naudin L, Broccardo C, Peterson KM, Koch C, Arnould I, Prades C, Duverger N, Funke H, Assman G, Dinger M, Dean M, Chimini G, Santamarina-Fojo S, Fredrickson DS, Denefle P, Brewer HB, Jr (1999) Human ATP-binding cassette transporter 1 (*ABC1*): genomic organization and identification of the genetic defect in the original Tangier disease kindred. Proc Natl Acad Sci USA 96:12685–12690

Rust S, Rosier M, Funke H, Real J, Amoura Z, Piette JC, Deleuze JF, Brewer HB, Duverger N, Denefle P, Assmann G (1999) Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. Nat Genet 22:352–355

Saito S, Iida A, Sekine A, Eguchi C, Miura Y, Nakamura Y (2001) Seventy genetic variations in human microsomal and soluble epoxide hydrolase (*EPHX1* and *EPHX2*) genes in the Japanese population. J Hum Genet 46:325–329

Santamarina-Fojo S, Peterson K, Knapper C, Qiu Y, Freeman L, Cheng JF, Osorio J, Remaley A, Yang XP, Haudenschild C, Prades C, Chimini G, Blackmon E, Francois T, Duverger N, Rubin EM, Rosier M, Denefle P, Fredrickson DS, Brewer HB, Jr (2000) Complete genomic sequence of the human *ABCA1* gene: analysis of the human and mouse ATP-binding cassette A promoter. Proc Natl Acad Sci USA 97:7987–7992

Sekine A, Saito S, Iida A, Mitsunobu Y, Higuchi S, Harigae S, Nakamura Y (2001) Identification of single nucleotide polymorphisms (SNPs) of human N-acetyltransferase genes NAT1, NAT2, AANAT, ARD1 and L1CAM in the Japanese population. J Hum Genet 46:314–319

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. (2001) The sequence of the human genome. Science 291:1304–1351