Table 2 Summary of *DPYD* SNPs detected in a Japanese population

SNP ID		Location	Position		Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Reported alleles	Allele frequency (341 subjects)
This study	dbSNP (NCBI)		NT_032977.7	From the translational initiation site or from the end of nearest exon				
MPJ6_DPD001[a]		5′-Flank	52206480	–609	TTGCTCGCCTCCC**C/T**TCCCCTCCCCGC			0.021
MPJ6_DPD002[a]		5′-Flank	52206348	–477	TTGAGGAGTTCC**T/G**GAAAATGCAGTT			0.026
MPJ6_DPD003[a]		5′-Flank	52206137	–266	CTCCCTCCCTCCC**/A**TTCTGCTTGCAG			0.045
MPJ6_DPD004[a]		5′-Flank	52206114	–243	AGGCTGGGGCGC**G/A**GAGCGGGCTGAA			0.0059
MPJ6_DPD005[a]		Exon 1	52205843	29	GTAAGGACTCGG**C/A**GGACATCGAGGT	Ala10Glu		0.0015
MPJ6_DPD006[g]	rs1801265	Exon 2	52168278	85	CATGCAACTCTG**T/C**GTTCCACTTCGG	Cys29Arg	*9	0.029
MPJ6_DPD007[a]		Intron 2	52168055	IVS2 + 158	TTTGAAAGTGTA**T/C**TTTTAATTACAC			0.0015
MPJ6_DPD008[a]		Intron 3	52113040	IVS3 + 23	GTCACCATAGCAA**/G**CAGTCACAGATG			0.0029
MPJ6_DPD009[a]		Exon 5	52006617	325	ATTTTGCAGAAC**T/A**ATTATGGAGCTG	Tyr109Asn		0.0029
MPJ6_DPD010[a]		Exon 5	52006491	451	GAGGGACCCATT**A/G**ATATTGGTGGAT	Asn151Asp		0.0088
MPJ6_DPD011[a]		Exon 5	52006468	474	ATTGCAGCAATT**T/C**GCTACTGAGGTA	Phe158Phe		0.0044
MPJ6_DPD012[a]		Intron 5	51984611	IVS5–115	CATATTAATACT**G/A**AAAATGTACTGC			0.021
MPJ6_DPD013[g]	rs2297595	Exon 6	51984484	496	GTATTCAAAGCAA**/G**TGAGTATCCCAC	Met166Val		0.022
MPJ6_DPD014[a]		Exon 6	51984341	639	GGGGTACTCTGA**C/T**ATCACTATATTT	Asp213Asp		0.0088
MPJ6_DPD015[a]		Exon 7	51976695	733	GTGAATTTTGAG**A/T**TTGAGCTAATGA	Ile245Phe		0.0015
MPJ6_DPD016[a]		Intron 7	51976602	IVS7 + 64	CTCTACACTAAA**G/T**ATTAACAGCAAA			0.0015
MPJ6_DPD017[a]		Exon 8	51964101	793	CTTTCAGTGAAT**G/A**AAATGACTCTTA	Glu265Lys		0.0015
MPJ6_DPD018[a]		Intron 8	51963953	IVS8 + 91	TTCAGACATTTT**C/T**GTGATGAAAGTT			0.0088
MPJ6_DPD019[a]		Intron 9	51878456	IVS9–120	TTTGATAGTGAC**A/T**CTTCATCCTGGA			0.0029
MPJ6_DPD020[b]		Exon 10	51878292	1003	ATACGGGGAGTC**G/T**TGATTGTACTTG	Val335Leu	*11	0.0015
MPJ6_DPD021[a]		Intron 10	51878143	IVS10 + 24	CCATCAGAAAAT**A/G**TGGAGTTGTACT			0.0015
MPJ6_DPD022[c]		Intron 10	51858934	IVS10-15	TTTCTTCTCTGTT**/C**CTGTTTTGTTTT			0.018
MPJ6_DPD023[a]		Intron 12	51800901	IVS12-11	AAGTATTGGTTT**G/A**TATTTTGCAGTC			0.038
MPJ6_DPD024[a]		Intron 12	51800899	IVS12-9	GTATTGGTTTGT**A/G**TTTTGCAGTCAC			0.0073
MPJ6_DPD025[a]		Exon 13	51800872	1543	TATGGAGCTTCC**G/A**TTTCTGCCAAGC	Val515Ile		0.0015
MPJ6_DPD026[a]		Exon 13	51800843	1572	ACTACCCCTCTTT**/G**TACACTCCTATT	Phe524Leu		0.0015
MPJ6_DPD027[g]	rs1801159	Exon 13	51800788	1627	GGATTGAAGTTT**A/G**TAAATCCTTTTG	Ile543Val	*5	0.283
MPJ6_DPD028[a]		Exon 13	51800749	1666	ACTCCAGCCACC**A/C**GCACATCAATGA	Ser556Arg		0.0015
MPJ6_DPD029	rs2786783	Intron 13	51800636	IVS13 + 39	AGAAATGTCTATC**/T**ATATATTTTAAT			0.283
MPJ6_DPD030	rs2811178	Intron 13	51800635	IVS13 + 40	GAAATGTCTATC**G/A**TATATTTTAATT			0.179
MPJ6_DPD031[a]		Intron 13	51735220_51735219	IVS13-47_-48	ATAAAGATTATA**–/TA**GCTTTTCTTTGT			0.0015
MPJ6_DPD032[a]		Exon 14	51735161	1752	GGACATTGTGAC**A/G**AATGTTTCCCCC	Thr584Thr		0.0015
MPJ6_DPD033[f]		Exon 14	51735139	1774	CCCAGAATCATCC**/T**GGGGAACCACCT	Arg592Trp		0.0015
MPJ6_DPD034[g]	rs17376848	Exon 14	51735017	1896	AAAGGCTGACTT**T/C**CCAGACAACGTA	Phe632Phe		0.139
MPJ6_DPD035[a]		Intron 14	51734989	IVS14 + 19	GTGATTTAACATC**/A**TAAAACAAGAGA			0.0088
MPJ6_DPD036[a]		Intron 14	51734908	IVS14 + 100	TTAATGTGTATA**T/G**TTTATTAAAGAA			0.0015
MPJ6_DPD037[a]		Intron 14	51667533	IVS14-123	GATTTATTTTTT**C/A**ACAGTTTGAAAA			0.155
MPJ6_DPD038[a]		Intron 14	51667431	IVS14-21	TGAACTTATATT**C/A**TTTTGTTTTTCT			0.0015
MPJ6_DPD039[d]		Intron 15	51667267	IVS15 + 75	TAAAGAGCTGCC**A/G**TGAGAAATAATA			0.155
MPJ6_DPD040[a]		Intron 16	51591373	IVS16-127	GGAATTTGAGAAA**/G**TATATCATGTAG			0.0015

Table 2 continued

SNP ID		Location	Position		Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Reported alleles	Allele frequency (341 subjects)
This study	dbSNP (NCBI)		NT_032977.7	From the translational initiation site or from the end of nearest exon				
MPJ6_DPD041[g]	rs7556439	Intron 16	51591340	IVS16-94	CAAGTTGGATTTG/TTCTTGCACGTCT			0.378
MPJ6_DPD042[a]		Intron 17	51591092	IVS17 + 34	GTTGCCCGCTATT/-GTAAATATTGGC			0.0015
MPJ6_DPD043[a]		Intron 17	51591079	IVS17 + 47	GTAAATATTGGCC/TCACATTATGTAG			0.0015
MPJ6_DPD044	rs1801160	Exon 18	51590313	2194	GGTGCCAATGGCG/ATTACAGCCACCA	Val732Ile	*6	0.015
MPJ6_DPD045[g]	rs12137711	Intron 18	51519982	IVS18-39	TATACTCAAGTGG/ATCAGTGTGCTAA			0.032
MPJ6_DPD046[c]		Exon 19	51519940	2303	TTTGTGTAGGGAC/AAGCAATCAGACC	Thr768Lys		0.028
MPJ6_DPD047[a]		Exon 19	51519819	2424	GTTTCTCCATAGT/CGGTGCTTCCGTC	Ser808Ser		0.0029
MPJ6_DPD048[a]		Exon 21	51383526	2678	TCATAGCAGAAAA/GCAAGATTAGACT	Asn893Ser		0.0029
MPJ6_DPD049[a]		Intron 21	51383358	IVS21 + 80	GTTTATTTACTGC/GTTAAATGTATCA			0.0015
MPJ6_DPD050[a]		Intron 21	51383325	IVS21 + 113	GTTTTAGAATTAT/AAATGAAAGTTTT			0.0015
MPJ6_DPD051[g]	rs11165777	Intron 21	51383302	IVS21 + 136	TTAAAAACATCTG/CTCCATGGTGAAA			0.0029
MPJ6_DPD052[a]		Intron 21	51383276	IVS21 + 162	CTGCATTTAAATT/GATAAAATAACCT			0.0029
MPJ6_DPD053[a]		Intron 22	51367150	IVS22 + 129	TTCTGCAACAGTA/GCATCTTTCTGTC			0.0073
MPJ6_DPD054[g]	rs290855	Intron 22	51364164	IVS22-69	GAGAAAAATGTTG/AACGCTAAAATGG			0.0029
MPJ6_DPD055[g]	rs17116357	Intron 22	51364153	IVS22-58	TAACGCTAAAATG/CGGGACATTGTTG			0.0029

[a] Novel variations detected in this study

[b] Kouwaki et al. 1998

[c] Collie-Duguid et al. 2000

[d] Seck et al. 2005

[e] Ogura et al. 2005

[f] Cho et al. 2007

[g] Variations overlapping with the HapMap project

In the 5′ flanking region, all four detected SNPs (-609C>T, -477T>G, -266C>A, -243G>A) were newly found at relatively high allele frequencies (0.006–0.05). However, these SNPs were not located near the proposed *cis*-regulatory promoter elements (Shestopal et al. 2000). The remaining 21 novel variations were found in intronic regions. Of these SNPs, IVS5–115G>A, IVS12–11G>A, and IVS14-123C>A were detected with allele frequencies of 0.021, 0.038, and 0.155, respectively, but others were rare (<0.01). They were not located in the exon-intron splicing junctions or branch sites.

Seventeen variations were already reported. The ID numbers in the dbSNP databases or references for these SNPs are described in Table 2. The well-known nonsynonymous SNPs, 1627A>G (**5*, Ile543Val), 2194G>A (**6*, Val732Ile), 85T>C (**9*, Cys29Arg), and 1003G>T (**11*, Val335Leu), were found in this study at allele frequencies of 0.283, 0.015, 0.029, and 0.0015, respectively. The allele frequencies of two reported SNPs, 496A>G (Met166Val) and 2303C>A (Thr768Lys), were 0.022 and 0.028, respectively. Recently, 1774C>T (Arg592Trp) was reported from a Korean population (Cho et al. 2007), and its allele frequency was 0.0015 in this study. Nine intronic variations, IVS10-15T>C, IVS13 + 39C>T, IVS13 + 40 G>A, IVS15 + 75A>G, IVS16-94G>T, IVS18-39G>A, IVS21 + 136G>C, IVS22-58G>C, and IVS22-69G>A, and one synonymous variation, 1896T>C (Phe632Phe), were found with various allele frequencies (0.003–0.378, Table 2). The variations previously detected in Japanese (Kouwaki et al. 1998; Yamaguchi et al. 2001; Ogura et al. 2005), 62G>A (Arg21Gln, **12*), 74G>A (His25Arg), 812delT (Leu271X), 1097G>C (Gly366Ala), 1156G>T (Glu386X, **12*), and 1714C>G (Leu572Val), were not found in our study. This might be due to their low frequencies.

Linkage disequilibrium (LD) analysis and haplotype block partition

LD analysis was performed by r^2 and |D′| using 18 SNPs (allele frequency ≥0.01) (Fig. 2). Strong linkages were observed in four pairs of SNPs: between -477T>G and 85T>C (Cys29Arg) (r^2 = 0.7025), between 496A>G (Met166Val) and IVS10-15T>C (r^2 = 0.7964), between 1627A>G (Ile543Val) and IVS13 + 39C>T (r^2 = 1.0), and between IVS14-123C>A and IVS15 + 75A>G (r^2 = 1.0). In addition, two known rare SNPs, IVS22-69G>A (rs290855) and IVS22-58G>C (rs17116357), were perfectly linked (r^2 = 1.0) (data not shown). As for |D′| values, only 43 pairs (28%) out of 153 pairs gave |D′| = 1.0, indicating that a number of recombinations had occurred within this gene. This is not surprising because *DPYD* is a huge gene of at least 950 kb in length with 3 kb of coding sequences. However, it was difficult to estimate past recombination events in *DPYD* from our data alone because our variations were mostly limited to exons and surrounding introns.

To define haplotype blocks, we utilized the HapMap data because SNPs were comprehensively genotyped with an average density of 1 SNP per 1.8 kb. Of 1,002 variations of *DPYD* genotyped by the HapMap project, 474 SNPs were polymorphic for 44 unrelated Japanese subjects. When the LD profiles for Japanese were obtained by Marker using the HapMap data, strong LD (|D′| > 0.75) clearly decays within introns 11, 12, 13, 14, 16, 18, and 20 (data not shown), suggesting that recombination had occurred in these regions. Based on these findings, the SNPs detected in our study were divided into six haplotype blocks (Figs. 1, 2). Block 1, the largest block, ranges from the 5′-untranslated region (5′-UTR) to intron 10 (347 kb), and includes 22 variations. Block 2 includes eight variations from IVS12-11G>A in intron 12 to IVS13 + 40G>A in intron 13. Block 3 includes six variations from IVS13-47_48insTA in intron 13 to IVS14 + 100T>G in intron 14. Block 4 contains only three SNPs, IVS14-123C>A, IVS14-21C>A and IVS15 + 75A>G, and ranges from intron 14 to intron 15. Block 5 consists of IVS16-94G>T and four rare variations from intron 16 to exon 18. Although the HapMap data showed a decline in LD in intron 20, we defined a block ranging from intron 18 to intron 22 as block 6 because only rare variations (allele frequencies <0.01) were detected downstream of intron 20 (exon 21, intron 21, and intron 22). The block partitioning based on the HapMap data fitted our SNPs well: more than 70% of SNP pairs in each block (block 1–6) gave pair-wise |D′| values greater than 0.8 (Fig. 2).

Haplotype estimation

Using 22, 8, 6, 3, 5, and 11 variations in blocks 1 to 6, 23 (block 1), 8 (block 2), 7 (block 3), 3 (block 4), 6 (block 5), and 11 (block 6) haplotypes were identified or inferred (Fig. 3). Probabilities of diplotype configurations in all six blocks were 100% for over 97% of the subjects. To discriminate our block haplotypes from the previously assigned alleles or haplotypes (*DPYD*1* to **13*), the mark, #, was used to indicate block haplotypes.

In block 1, the most dominant haplotype without any variation was #*1a* (0.818 in frequency), followed by #*1b* (0.045), #*9c* (0.022), and #*1c* (0.021). As suggested by LD (Fig. 2), #*9c*, the major subtype of the #*9* group bearing 85T>C (Cys29Arg), also harbored -477T>G in the 5′-UTR. Known nonsynonymous SNP, 496A>G (Met166Val), was assigned to three haplotypes, #*9d*, #*166Va*, and #*166Vb*.

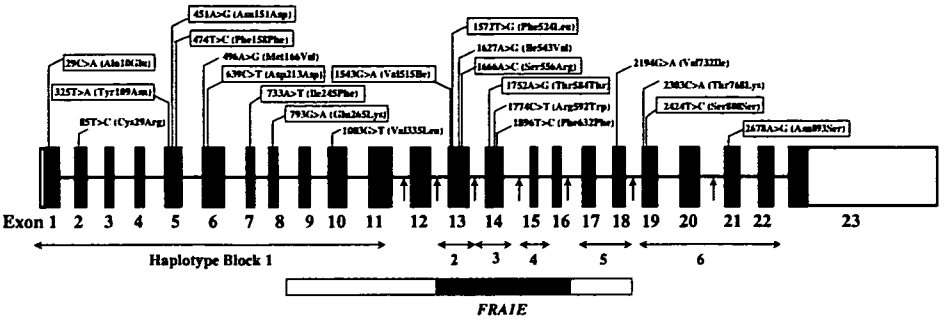


Fig. 1 Twenty-one variations detected in the coding exons are depicted in the schematic diagram of the *DPYD* gene. Fourteen novel variations are enclosed by *squares*. The recombination spots were estimated based on the LD profiles obtained from Japanese data in the HapMap project and indicated by *arrows*. The borders (between introns 8 and 18 of the *DPYD*) and core region (between introns 12 and 16) of *FRA1E* identified by Hormozian et al. (2007) are indicated as an *open* and *closed box*, respectively

In block 2, four haplotypes, #*1a* (0.529), #*5a* (0.245), #*1b* (0.176), and #*5b* (0.038), were major in Japanese and accounted for 99% of all inferred haplotypes. Two subtypes of the #*5* group, #*5a* and #*5b*, both of which harbored Ile543Val (**5*) and IVS13 + 39C>T, were distinguished by a novel intronic SNP, IVS12-11G>A.

As for block 3, in addition to #*1a* (0.848), #*1b* harboring the synonymous SNP, 1896T>C (Phe632Phe), was found at a relatively high frequency (0.138).

Block 4 is simple and comprises only three haplotypes, #*1a* (0.845), #*1b* (0.154) and #*1c* (0.0015). The second frequent haplotype, #*1b*, harbored perfectly linked SNPs, IVS14-123C>A and IVS15 + 75A>G.

Block 5 contained IVS16-94G>T, the most frequent SNP among the 55 SNPs found in this study, which was assigned to #*1b* with a frequency of 0.374. This block also contained the known nonsynonymous SNP, 2194G>A (Val732Ile, **6*), which was assigned to #*6a* (0.015).

In block 6, the most dominant haplotype was #*1a* (0.915). It was followed by #*1b* (0.032) with IVS18-39G>A and #*768K* (0.028) with 2303C>A (Thr768Lys).

The HapMap data include nine SNPs that we detected (Table 2). Of them, six, 85T>C (rs1801265), 496A>G (rs2297595), 1627A>G (rs1801159), 1896T>C (rs17376848), IVS16-94G>T (rs7556439) and IVS18-39G>A (rs12137711), were suitable for haplotype tagging SNPs (htSNPs) to capture the block haplotypes, block 1 #*9*, block 1 #*166V*, block 2 #*5*, block 3 #*1b*, block 5 #*1b*, and block 6 #*1b*, respectively. IVS21 + 136G>C (rs11165777) and IVS22-69G>A (rs290855)/IVS22-58G>C (rs1711 6357), were the marker SNPs for block 6 #*1e* and #*1f*, respectively, but very rare (allele frequencies = 0.003) in Japanese. The six SNPs, especially 85T>C (rs1801265) and 496A>G (rs2297595), were in strong LD ($r^2 > 0.8$) with other HapMap SNPs in Japanese (Table 3), indicating that many HapMap SNPs were concurrently linked on the same haplotypes.

Next, the combinations of block haplotypes (inter-block haplotypes) were analyzed focusing on the haplotypes with frequencies of >0.01 in each block (Fig. 4). Between blocks 1 and 2, both #*1a* and #*1b* in block 1 were complicatedly associated with various haplotypes in block 2. It should be noted that #*9c* in block 1 was linked either with block 2 #*1b* (0.016 in absolute frequency) or with block 2 #*5a* (0.006, not shown in Fig. 4). #*1c* in block 1 was completely linked with block 2 #*1a*. #*151D* in block 1 (not shown in Fig. 4), which was a rare haplotype (0.009) harboring 451A>G (Asn151Asp), was completely linked with #*5a* in block 2.

Between blocks 2 and 3, both #*5b* and #*1b* in block 2 were mostly linked with #*1a* in block 3, whereas both #*1a* and #*5a* in block 2 were complicatedly linked with #*1a*, #*1b*, or other rare haplotypes such as #*1c* (not shown in Fig. 4) in block 3. Between blocks 3 and 4 and between blocks 4 and 5, no strong associations of block haplotypes were observed except for the linkage of block 5 #*6a* to block 4 #*1a*. Between blocks 5 and 6, most of #*1b* and all of #*6a* in block 5 were linked with #*1a* in block 6. Although #*1a* in block 6 was associated with various haplotypes in block 5, #*1b* in block 6 was completely linked with #*1a* in block 5.

Among the six blocks, the following combinations were major: #*1a* (block 1)–#*1a* (block 2) –#*1a* (block 3)–#*1a* (block 4)–#*1a* (block 5)–#*1a* (block 6) (0.239 in frequency), #*1a*–#*5a*–#*1a*–#*1a*–#*1b*–#*1a* (0.081), #*1a*–#*1a*–#*1a*–#*1a* –#*1b*–#*1a* (0.075), #*1a*–#*5a*–#*1a*–#*1a*–#*1a*–#*1a* (0.070), #*1a*–#*1b*–#*1a*–#*1a*–#*1a*–#*1a* (0.060) and #*1a*–#*1a*–#*1b*–#*1a*–#*1a*–#*1a* (0.051).

Ethnic differences in distributions of *DPYD* SNPs and haplotypes

We compared SNP and haplotype distributions in Japanese with those in other ethnic groups reported in the literature

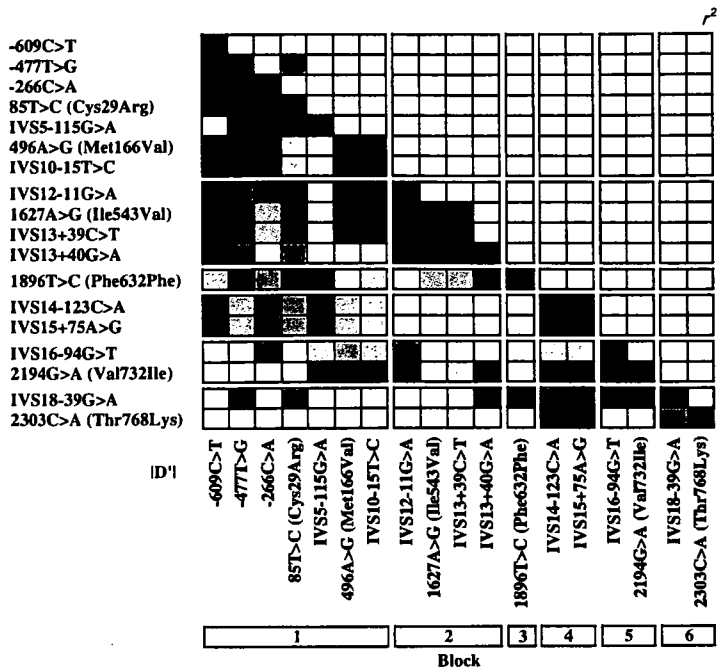


Fig. 2 Linkage disequilibrium (*LD*) analysis of *DPYD*. Pairwise LD between 18 common SNPs (>0.01 in allele frequencies) is expressed as r^2 (*upper*) and |D'| (*lower*) by a 10-graded *blue color*. The *denser color* indicates higher linkage. The haplotype block partition based on LD measure |D'| of HapMap data in Japanese is also indicated

or HapMap project. Notably, IVS14 + 1G>A (*2), 1897delC (Pro633GlnfsX5, *3), 1601G>A (Ser534Asn, *4), 295_298delTCAT (Phe100SerfsX15, *7), 703C>T (Arg235Trp, *8), 2983G>T (Val995Phe, *10), 62G>A (Arg21Gln, *12), 1156G>T (Glu386X, *12), and 1679T>G (Ile560Ser, *13) were not found in this study. Furthermore, several SNPs showed marked differences in allele frequencies among Japanese and other ethnic groups (Table 4).

The allele frequency of 85T>C (Cys29Arg, *9), the tagging SNP for block 1 #9, was quite different between Asians and Caucasians. Its allele frequency in Japanese (0.029 in this study) and Taiwanese (0.022) (Hsiao et al. 2004) was much lower than that in Caucasians (0.185–0.194) (Seck et al. 2005; Morel et al. 2006).

The SNP 496A>G (Met166Val) in block 1 is found at a lower allele frequency in Japanese (0.022) than in Caucasians (0.080) (Seck et al. 2005). Seck et al. (2005) inferred two haplotypes harboring 496A>G (Met166Val) from 157 Caucasians: *hap5* (#9d in this study) harboring additional 85T>C (Cys29Arg) and IVS10-15T>C and *hap11* concurrently harboring IVS10-15T>C alone with frequencies of 0.040 and 0.014, respectively. In our haplotype analysis, #166Va (0.012) corresponding to *hap11* (0.014) was found with a similar frequency in Japanese, whereas the frequency of #9d (0.006) was much lower than that of the corresponding haplotype, *hap5* (0.040) in Caucasians.

1627A>G (Ile543Val, *5) in block 2 was found with comparable allele frequencies among Japanese (0.283 in this study), Caucasians (0.14-0.275) (Seck et al. 2005; Ridge et al. 1998a), African-Americans (0.227) (Wei et al. 1998), and Taiwanese (0.210–0.283) (Wei et al. 1998; Hsiao et al. 2004).

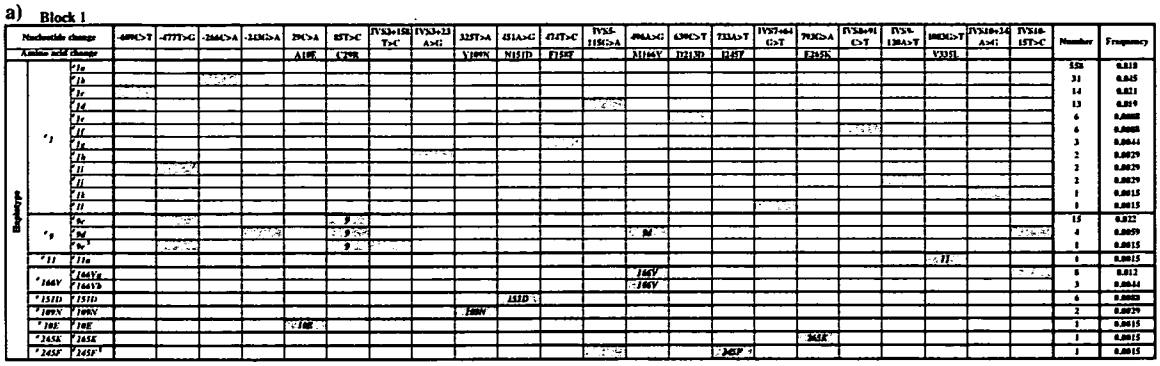
The allele frequency (0.015) of 2194G>A (Val732Ile, *6) in block 5 in our Japanese population is slightly lower than that previously reported in Caucasians (0.022-0.058) (Seck et al. 2005; Ridge et al. 1998a) and Finish (0.067) (Wei et al. 1998), but is comparable to that in Taiwanese (0.012-0.014) (Wei et al. 1998; Hsiao et al. 2004) and African-Americans (0.019) (Wei et al. 1998).

Ethnic differences in the allele frequencies were also observed with synonymous and intronic variations (Table 4). The allele frequency of 1896T>C (Phe632Phe), which tags block 3 #1b, was higher in Japanese (0.139 in this study) than in Caucasians (0.035) (Seck et al. 2005). *Hap13* assigned in 157 Caucasians by Seck et al. (2005) is the counterpart of block 3 #1b, and its frequency (0.012) was much lower than that in Japanese (0.138).

In contrast, IVS10-15T>C linked to 85T>C (*9) or 496A>G (#166V) within block 1 showed a lower allele frequency in Japanese (0.018) than in Caucasians (0.127). Seck et al. (2005) assigned *hap7* as the haplotype containing IVS10-15T>C alone with a haplotype frequency of 0.03 in Caucasians. In Japanese, however, the corresponding haplotype was not found.

Allele frequencies of IVS18-39G>A and IVS22-69G>A, which are tagging SNPs for block 6 #1b and #1f, respectively, are lower in Japanese (0.032 and 0.003, respectively) than in Caucasians (0.105 and 0.183, respectively).

a) Block 1



b) Block 2

Nucleotide change			IVS12-11G>A	IVS12-9A>G	1543G>A	1572T>G	1627A>G	1666A>C	IVS13+39 C>T	IVS13+40 G>A	Number	Frequency
Amino acid change					V515I	F524L	I543V	S556R				
Haplotype	*1	*1a									361	0.529
		*1b									120	0.176
		*1c									5	0.0073
	*5	*5a					5				167	0.245
		*5b					5				26	0.038
	*515I	*515I			515I						1	0.0015
	*524L	*524L				524L					1	0.0015
	*556R	*556R§						556R			1	0.0015

c) Block 3

Nucleotide change			IVS13-47_48insTA	1752A>G	1774C>T	1896T>C	IVS14+19 C>A	IVS14+100 T>G	Number	Frequency
Amino acid change				T584T	R592W	F632F				
Haplotype	*1	*1a							578	0.848
		*1b							94	0.138
		*1c							6	0.0088
		*1d							1	0.0015
		*1e							1	0.0015
		*1f§							1	0.0015
	*592W	*592W			592W				1	0.0015

d) Block 4

Nucleotide change			IVS14-123C>A	IVS14-21C>A	IVS15+75 A>G	Number	Frequency
Amino acid change							
Haplotype	*1	*1a				576	0.845
		*1b				105	0.154
		*1c§				1	0.0015

e) Block 5

Nucleotide change			IVS16-127A>G	IVS16-94G>T	IVS17+34del T	IVS17+47C>T	2194G>A	Number	Frequency
Amino acid change							V732I		
Haplotype	*1	*1a						414	0.607
		*1b						255	0.374
		*1c§						1	0.0015
		*1d§						1	0.0015
		*1e§						1	0.0015
	*6	*6a					6	10	0.015

f) Block 6

Nucleotide change			IVS18-39G>A	2303C>A	2424T>C	2678A>G	IVS21+80 C>G	IVS21+113 T>A	IVS21+136 G>C	IVS21+162 T>G	IVS22+129 A>G	IVS22-69 G>A	IVS22-58 G>C	Number	Frequency
Amino acid change				T768K	S808S	N893S									
Haplotype	*1	*1a												624	0.915
		*1b												22	0.032
		*1c												5	0.0073
		*1d												2	0.0029
		*1e												2	0.0029
		*1f												2	0.0029
		*1g												2	0.0029
		*1h												1	0.0015
		*1i												1	0.0015
	*768K	*768K		768K										19	0.028
	*893S	*893S				893S								2	0.0029

Fig. 3 Block haplotypes in *DPYD* of block 1 (**a**), block 2 (**b**), block 3 (**c**), block 4 (**d**), block 5 (**e**), and block 6 (**f**) in a Japanese population. The nucleotide positions were numbered based on the cDNA sequence (A of the translational start codon is +1) or from the nearest exon. *White cell* wild-type, *gray cell* nucleotide alteration. §The haplotypes were inferred in only one patient and ambiguous except for marker SNPs

Table 3 Linkages of haplotype-tagging SNPs with HapMap SNPs for *DPYD*

Haplotype-tagging SNPs in *DPYD*	dbSNP ID (NCBI)	Block haplotype in this paper	HapMap SNPs with close linkages ($r^2 > 0.8$)[a]
85T>C (Cys29Arg)	rs1801265	Block 1 #*9*	rs10747488, rs7526108, rs4421623, rs4379706, rs4523551,rs11165921, rs9661794, rs6677116, rs6604093, rs17379561, rs10747491, rs10747492, rs12062845, rs7524038, rs10875112, rs4394693, rs10875113, rs4970722, rs9727548, rs10875118, rs9662719, rs12077442, rs4394694, rs9727976, rs4246515, rs6692580
496A>G (Met166Val)	rs2297595	Block 1 #*166V*	rs2786543, rs2811215, rs2811214, rs2786544, rs2248658, rs11165897, rs2786490, rs2811203, rs2811202, rs2811200, rs2811198, rs2786503, rs2811196, rs2786505, rs2811195, rs2811194, rs12073839, rs6663670, rs7512910, rs2151563, rs2786509, rs3790387, rs3790389
1627A>G (Ile543Val)	rs1801159	Block 2 #*5*	rs1415682, rs952501, rs2811187, rs2786778, rs2786774, rs2811183, rs17116806, rs2786780, rs1801159, rs2786771, rs2297780, rs2297779, rs12729863
1896T>C (Phe632Phe)	rs7556439	Block 3 #*1b*	rs12073650
IVS16-94G>T	rs7556439	Block 5 #*1b*	rs693680, rs827500, rs499009, rs7518848, rs553388, rs507170, rs628959, rs991544, rs526645, rs1609519
IVS18-39G>A	rs12137711	Block 6 #*1b*	rs12120068, rs12116905

[a] All SNPs are in the same block

Taken together, our data demonstrated considerable differences in the haplotype distributions in blocks 1, 3 and 6 between Japanese and Caucasians.

Discussion

This study provides Japanese data on the genetic variations of *DPYD*, a gene encoding a key enzyme catalyzing degradation of the well-known anticancer drug 5-FU. Nine novel (Ala10Glu, Tyr109Asn, Asn151Asp, Ile245Phe, Glu265Lys, Val515Ile, Phe524Leu, Ser556Arg, and Asn893Ser) and seven known nonsynonymous variations (Cys29Arg, Met166Val, Val335Leu, Ile543Val, Arg592Trp, Val732Ile, and Thr768Lys) were found in our Japanese population (Table 2 and Fig. 1). The association analysis between the genotypes and 5-FU pharmacodynamics is now on-going.

Uneven distributions of coding SNPs over 23 *DPYD* exons were pointed out in the previous review by van Kuilenburg (2004). The author indicated that 81% of all reported variations were confined to exons 2–14, representing 61% of the coding sequences, and typical hotspots of variation were localized in exons 2, 6, and 13. Our Japanese data also revealed that 17 out of 21 coding variations (81%) were localized in exons 1–14, and that more than three variations were detected in exons 5, 13, and 14 (Fig. 1). Recently, Hormozian et al. (2007) have reported that the common chromosomal fragile site on 1p21.2, *FRA1E*, spans 370 kb of genomic sequence between introns 8 and 18 of *DPYD*, and that its core region with the highest fragility is located between introns 12 and 16. The instability at the core of *FRA1E* might be associated with the high mutational rates and recombinogenic nature from intron 12 to 14 of *DPYD* (Fig. 1).

To estimate potential functional consequences of the amino acid substitutions, we examined whether the positions of amino acid changes are located in highly conserved areas or potentially critical regions of the molecule (for example, substrate recognition sites or binding regions of prosthetic groups). We also considered the locations of the residues in a three-dimensional (3D) framework provided by the crystal structures of pig DPD, which have recently been determined in complexes with NADPH and substrate (5-FU) (Dobritzsch et al. 2001) or inhibitors (Dobritzsch et al. 2002). The amino acid sequences of pig and human DPD are 93% identical (Mattison et al. 2002), and the substituted residues and their neighboring residues are conserved between both enzymes. From these points of view, it is speculated that at least two substitutions (Glu265Lys and Arg592Trp) might impact the structure and function of DPD as discussed below.

Glu265 is located on the loop following to the third β sheet (IIβ3) in the FAD binding domain II (Dobritzsch et al. 2001). Glu265 is conserved among four mammalian species (human, mouse, rat, and pig), although it is replaced with aspartic acid in bovine and *Drosophila melanogaster* DPDs (Mattison et al. 2002). In the 3D structure of pig DPD (Fig. 5a), Glu265 is in close proximity to Lys259. The substitution, Lys259Glu, was

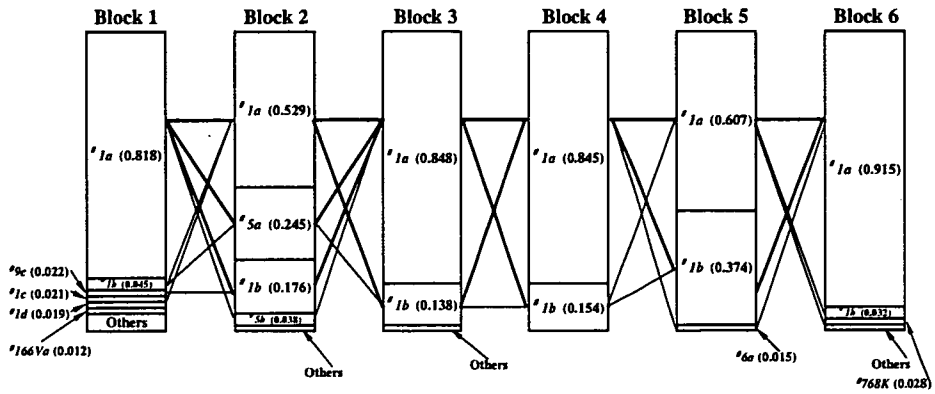


Fig. 4 The combinations of block haplotypes in Japanese. *Thick lines* represent combinations with frequencies over 10%, and *thin lines* represent combinations with frequencies of 1.0–9.9%

detected in the patient exhibiting severe mucositis during cyclophosphamide/methotrexate/5-FU chemotherapy (Gross et al. 2003). Furthermore, the adjacent Leu261 interacts via the main chain atoms with the N6, N1, and N3 atoms of adenine of FAD, and has an important role in the proper orientation of the adenine moiety in the FAD-binding pocket (Dobritzsch et al. 2001). Moreover, the carboxyl group (Glu265-Oε)might form hydrogen bonds to the main chain nitrogen of Ser260 next to Leu261. Thus, the change in polarity from negative to positive by the novel Glu265Lys substitution is likely to cause structural changes affecting proper binding of FAD.

Arg592 is located at one (IVβc) of the additional four-stranded antiparallel β sheets (IVβc-βf) inserted at the top of a typical $(\alpha/\beta)_8$ barrel fold in the FMN-binding domain IV (Dobritzsch et al. 2001). Arg592 is completely conserved among the above-mentioned six species (Mattison et al. 2002), suggesting its functional importance. Arg592 closely contacts Met599 (2.9 Å) and Gln604 (2.8 Å) in the same subunit and Ser994 (2.9 Å) in another subunit (Fig. 5B). The substitution of tryptophan for Arg592 is likely to weaken these interactions due to altered hydrophobicity and electrostatic changes. Arg592Trp was recently reported from a Korean population with an allele frequency of 0.004, although its functional significance remains to be confirmed(Cho et al. 2007).

As for known *DPYD* alleles, their distributions in several populations are becoming more evident by recent reports. For example, IVS14 + 1G>A (*2) (van Kuilenburg 2004), 295_298delTCAT (Phe100SerfsX15, *7) (Seck et al. 2005), 1679T>G (Ile560Ser, *13) (Collie-Duguid et al. 2000; Morel et al. 2006) 2846A>T (Asp949Val) (Seck et al. 2005; Morel et al. 2006), all of which are associated with decreased DPD activities, are detected in Caucasians with allele frequencies of 0.01–0.02, 0.003, 0.001 and 0.006–0.008, respectively. However, none of them were detected in our Japanese samples, while 1003G>T (Val335Leu, *11) and 2303C>A (Thr768Lys) have been found only in Japanese, indicating that variations with clinical relevance do not overlap between Caucasians and Japanese.

2303C>A (Thr768Lys), which was originally found in a Japanese female volunteer with very low DPD activity (Ogura et al. 2005), is relatively frequent in Japanese (allele frequency = 0.0279). Functional characterization in vitro revealed that 768Lys caused thermal instability of the variant protein without changing its affinity for NADPH or kinetic parameters toward 5-FU. Therefore, they might cause 5-FU-related toxicities in Japanese.

1003G>T (Val335Leu, *11) was found in a Japanese family with decreased DPD activity by Kouwaki et al. (1998). By in vitro expression in *E. coli*, they demonstrated that the variant protein with Leu335 showed a significant loss of activity (about 17% of the wild-type protein). Dobritzsch et al. (2001) suggested from the 3D structure of pig DPD that Val335Leu, in spite of a conservative change, disturbs packing interactions in the hydrophobic core formed by IIIβ3 and IIIα3 within the Rossman-motif, thereby affecting NADPH binding. In our study, heterozygous 1003G>T (Val335Leu) was found from a patient administrated 5-FU (allele frequency = 0.0015), who also has seven other variations: IVS12–11G>A, 1896T>C (Phe632Phe), and IVS16–94G>T are heterozygous, and 1627A>G (Ile543Val), IVS13 + 39C>T, IVS14–123C>A, and IVS15 + 75A>G are homozygous, indicating that at least Val335Leu is linked to Ile543Val (*5).

On the other hand, Caucasians and Japanese share four variations: *5 (Ile543Val), *9 (Cys29Arg), Met166Val, and *6 (Val732Ile), although their allele frequencies were different, especially for *9 (Table 4). Because they have not necessarily correlated with phenotypic changes (e.g., differences in DPD enzyme activity, 5-FU pharmacokinetics and pharmacodynamics) (Collie-Duguid et al. 2000; Johnson et al. 2002; Zhu et al. 2004; Seck et al. 2005; Ridge et al. 1998a, 1998b; Hsiao et al. 2004), all of these variations are generally accepted as common polymorphisms that result in unaltered function. Consistent with this, van Kuilenburg et al. (2002) suggested that the

Table 4 Allele frequencies of common *DPYD* SNPs in different populations

Nucleotide change (amino acid change)	Allele or tagged haplotypes	Population	Allele frequency	Number of subjects	Reference
85T>C	*9	Caucasian	0.194	157	Seck et al. 2005
(Cys29Arg)	(Block 1 #9)	French Caucasian	0.185	487	Morel et al. 2006
		Japanese	0.037	107	Yamaguchi et al. 2001
		Japanese	0.029	341	This study
		Taiwanese	0.022	300	Hsiao et al. 2004
496A>G	Block 1 #166V	Caucasian	0.080	157	Seck et al. 2005
(Met166Val)		Japanese	0.022	341	This study
IVS10-15T>C	Block 1 #166Va, #9d	Caucasian	0.127	157	Seck et al. 2005
		Japanese	0.018	341	This study
1627A>G	*5	Caucasian	0.140	157	Seck et al. 2005
(Ile543Val)	(Block 2 #5)	Caucasian	0.275	60	Ridge et al. 1998a
		Finnish	0.072	90	Wei et al. 1998
		African-American	0.227	105	Wei et al. 1998
		Japanese	0.352	50	Wei et al. 1998
		Japanese	0.283	341	This study
		Taiwanese	0.210	131	Wei et al. 1998
		Taiwanese	0.283	300	Hsiao et al. 2004
1896T>C	Block 3 #1b	Caucasian	0.035	157	Seck et al. 2005
(Phe632Phe)		Japanese	0.098	107	Yamaguchi et al. 2001
		Japanese	0.139	341	This study
		Han Chinese	0.133	45	HapMap
IVS15 + 75A>G	Block 4 #1b	Caucasian	0.166	157	Seck et al. 2005
		Japanese	0.155	341	This study
IVS16-94G>T	Block 5 #1b	Caucasian	0.415	59	HapMap
		Yorba	ND	60	HapMap
		Japanese	0.455	44	HapMap
		Japanese	0.378	341	This study
		Han Chinese	0.333	45	HapMap
2194G>A	*6	Caucasian	0.022	157	Seck et al. 2005
(Val732Ile)	(Block 5 #6)	Caucasian	0.058	60	Ridge et al. 1998a
		Finnish	0.067	90	Wei et al. 1998
		African-American	0.019	105	Wei et al. 1998
		Japanese	0.044	50	Wei et al. 1998
		Japanese	0.015	341	This study
		Taiwanese	0.014	131	Wei et al. 1998
		Taiwanese	0.012	300	Hsiao et al. 2004
IVS18-39G>A	Block 6 #1b	Caucasian	0.105	157	Seck et al. 2005
		Caucasian	0.100	60	HapMap
		Yorba	0.017	60	HapMap
		Japanese	0.044	45	HapMap
		Japanese	0.032	341	This study
		Han Chinese	0.022	45	HapMap
IVS22-69G>A	Block 6 #1f	Caucasian	0.183	60	HapMap
		Yorba	0.400	60	HapMap
		Japanese	ND	45	HapMap
		Japanese	0.003	341	This study
		Han Chinese	ND	45	HapMap

ND not detected

substitution Cys29Arg on the protein surface was unlikely to alter DPD activity. However, conflicting results were reported regarding *9 (Vreken et al. 1997, van Kuilenburg et al. 2000), *6 (van Kuilenburg et al. 2000), and Met166Val (van Kuilenburg et al. 2000; Gross et al. 2003). To interpret these inconsistencies, haplotype analysis of *DPYD* might be helpful. Especially for *9 and Met166Val in Japanese, functional involvement of –477T>G (block 1 #9c and #9e), –243G>A (block 1 #9d), IVS10-15T>C (block 1 #9d and #166Va) and many other HapMap SNPs linked to *9 and Met166Val (Table 3) needs clarification.

The HapMap project provides genotype data of more than 1,000 sites located mostly in the intronic regions of *DPYD* for four different populations (Nigerian, Chinese,

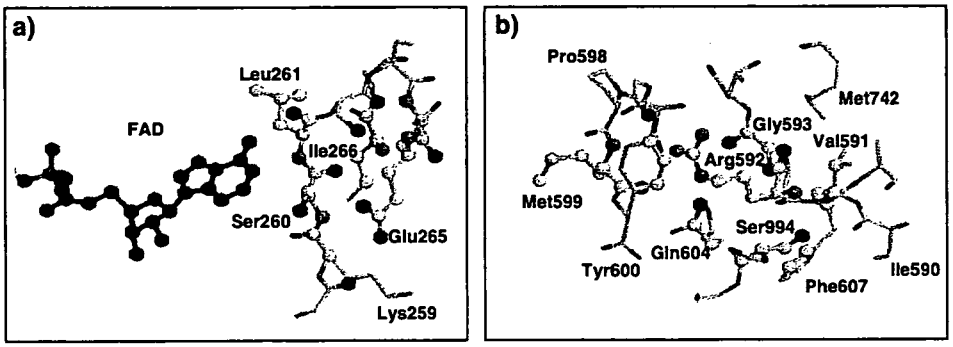


Fig. 5 Stereo view of the variation sites in pig DPD (accession code of the Protein Data Bank: 1gth). Glu265 (a), Arg592 (b) and their adjacent residues are shown as *ball-and-stick* models with oxygens in *red*, nitrogens in *blue*, carbons in *gray* and sulfur in *yellow*. The adenosine moiety of the cofactor FAD is also shown in *pink* (a)

Japanese and Caucasians). HapMap data on 44 unrelated Japanese subjects showed that 476 variations are polymorphic, whereas 529 are monomorphic, and the average density of polymorphic markers is 1 SNP per 1,772 bp. In contrast, our study focused on exons and surrounding introns to detect variations, and only nine variations overlapped with the HapMap data. Therefore, we could not utilize the HapMap data to further identify common subtypes of *#1* to be discriminated by many intronic HapMap SNPs in each block. However, most of the frequent SNPs are unlikely to be associated with substantially decreased DPD activity because DPD activity in the healthy Japanese population ($N = 150$) showed a unimodal Gaussian distribution (Ogura et al. 2005).

On the other hand, in 60 unrelated Caucasian subjects in the HapMap project, 617 are polymorphic, whereas 383 are monomorphic. LD profiles of these polymorphisms were compared between Caucasians and Japanese by using the program Marker (http://www.gmap.net/marker). Strong LD ($|D'| > 0.75$) clearly decays within introns 11, 12, 13, 14, 16, 18, and 20 in Japanese, whereas, similar decays are observed within introns 13, 14, 18, and 20, but are not obvious within introns 11, 12, and 16 in Caucasians (data not shown). Moreover, strong LD decays within intron 3 in Caucasians. Therefore, the LD blocks are considerably different between Japanese and Caucasians. Along with the marked differences in allele frequencies of several variations (Table 4), these results suggest that the haplotype structures in *DPYD* are quite different between the two populations.

In conclusion, we found 55 variations, including 38 novel ones, in *DPYD* from 341 Japanese subjects. Nine novel nonsynonymous SNPs were found, some of which were assumed to have impact on the structure and function of DPD. As for known variations, we obtained their accurate allele frequencies in a Japanese population of a large size and showed that variations with clinical relevance do not overlap between Caucasians and Japanese. In Japanese, 2303C>A (Thr768Lys) and 1003G>T (Val335Leu) might play important roles in 5-FU-related toxicity. Along with differences in haplotype structures between Japanese and Caucasians, these findings suggest that ethnic-specific tagging SNPs should be considered on genotyping *DPYD*. Thus, the present information would be useful for pharmacogenetic studies for evaluating the efficacy and toxicity of 5-FU in Japanese and probably in East Asians.

Acknowledgments We thank Ms. Chie Sudo for her secretarial assistance. This study was supported in part by the Program for the Promotion of Fundamental Studies in Health Sciences (05–25) of the National Institute of Biomedical Innovation and in part by the Health and Labor Sciences Research Grants from the Ministry of Health, Labor and Welfare.

References

Bakkeren JA, De Abreu RA, Sengers RC, Gabreels FJ, Maas JM, Renier WO (1984) Elevated urine, blood and cerebrospinal fluid levels of uracil and thymine in a child with dihydrothymine dehydrogenase deficiency. Clin Chim Acta 140:247–256

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Cho HJ, Park YS, Kang WK, Kim JW, Lee SY (2007) Thymidylate synthase (TYMS) and dihydropyrimidine dehydrogenase (DPYD) polymorphisms in the Korean population for prediction of 5-fluorouracil-associated toxicity. Ther Drug Monit 29:190–196

Collie-Duguid ES, Etienne MC, Milano G, McLeod HL (2000) Known variant DPYD alleles do not explain DPD deficiency in cancer patients. Pharmacogenetics 10:217–223

Dobritzsch D, Schneider G, Schnackerz KD, Lindqvist Y (2001) Crystal structure of dihydropyrimidine dehydrogenase, a major determinant of the pharmacokinetics of the anti-cancer drug 5-fluorouracil. Embo J 20:650–660

Dobritzsch D, Ricagno S, Schneider G, Schnackerz KD, Lindqvist Y (2002) Crystal structure of the productive ternary complex of dihydropyrimidine dehydrogenase with NADPH and 5-iodouracil. Implications for mechanism of inhibition and electron transfer. J Biol Chem 277:13155–13166

Etienne MC, Lagrange JL, Dassonville O, Fleming R, Thyss A, Renee N, Schneider M, Demard F, Milano G (1994) Population study of dihydropyrimidine dehydrogenase in cancer patients. J Clin Oncol 12:2248–2253

Gross E, Ullrich T, Seck K, Mueller V, de Wit M, von Schilling C, Meindl A, Schmitt M, Kiechle M (2003) Detailed analysis of five mutations in dihydropyrimidine dehydrogenase detected in

cancer patients with 5-fluorouracil-related side effects. Hum Mutat 22:498

Grem JL (1996) Fluoropyrimidines. In: Chabner BA, Longo DL (eds) Cancer chemotherapy and biotherapy, 2nd edn. Lippincott-Raven, Philadelphia, pp 149–197

Heggie GD, Sommadossi JP, Cross DS, Huster WJ, Diasio RB (1987) Clinical pharmacokinetics of 5-fluorouracil and its metabolites in plasma, urine, and bile. Cancer Res 47:2203–2206

Hormozian F, Schmitt JG, Sagulenko E, Schwab M, Savelyeva L (2007) *FRA1E* common fragile site breaks map within a 370 kilobase pair region and disrupt the dihydropyrimidine dehydrogenase gene (*DPYD*). Cancer Lett 246:82–91

Hsiao HH, Yang MY, Chang JG, Liu YC, Liu TC, Chang CS, Chen TP, Lin SF (2004) Dihydropyrimidine dehydrogenase pharmacogenetics in the Taiwanese population. Cancer Chemother Pharmacol 53:445–451

Johnson MR, Wang K, Diasio RB (2002) Profound dihydropyrimidine dehydrogenase deficiency resulting from a novel compound heterozygote genotype. Clin Cancer Res 8:768–774

Kitamura Y, Moriguchi M, Kaneko H, Morisaki H, Morisaki T, Toyama K, Kamatani N (2002) Determination of probability distribution of diplotype configuration (diplotype distribution) for each subject from genotypic data using the EM algorithm. Ann Hum Genet 66: 183–193

Kouwaki M, Hamajima N, Sumi S, Nonaka M, Sasaki M, Dobashi K, Kidouchi K, Togari H, Wada Y (1998) Identification of novel mutations in the dihydropyrimidine dehydrogenase gene in a Japanese patient with 5-fluorouracil toxicity. Clin Cancer Res 4:2999–3004

Lu Z, Zhang R, Diasio RB (1993) Dihydropyrimidine dehydrogenase activity in human peripheral blood mononuclear cells and liver: population characteristics, newly identified deficient patients, and clinical implication in 5-fluorouracil chemotherapy. Cancer Res 53:5433–5438

Lu Z, Zhang R, Carpenter JT, Diasio RB (1998) Decreased dihydropyrimidine dehydrogenase activity in a population of patients with breast cancer: implication for 5-fluorouracil-based chemotherapy. Clin Cancer Res 4:325–329

Martz E (2002) Protein explorer: easy yet powerful macromolecular visualization. Trends Biochem Sci 27:107–109

Mattison LK, Johnson MR, Diasio RB (2002) A comparative analysis of translated dihydropyrimidine dehydrogenase cDNA; conservation of functional domains and relevance to genetic polymorphisms. Pharmacogenetics 12:133–144

McLeod HL, Collie-Duguid ES, Vreken P, Johnson MR, Wei X, Sapone A, Diasio RB, Fernandez-Salguero P, van Kuilenberg AB, van Gennip AH, Gonzalez FJ (1998) Nomenclature for human DPYD alleles. Pharmacogenetics 8:455–459

Morel A, Boisdron-Celle M, Fey L, Soulie P, Craipeau MC, Traore S, Gamelin E (2006) Clinical relevance of different dihydropyrimidine dehydrogenase gene single nucleotide polymorphisms on 5-fluorouracil tolerance. Mol Cancer Ther 5:2895–2904

Naguib FN, el Kouni MH, Cha S (1985) Enzymes of uracil catabolism in normal and neoplastic human tissues. Cancer Res 45:5405–5412

Nishiyama T, Ogura K, Okuda H, Suda K, Kato A, Watabe T (2000) Mechanism-based inactivation of human dihydropyrimidine dehydrogenase by (E)-5-(2-bromovinyl)uracil in the presence of NADPH. Mol Pharmacol 57:899–905

Ogura K, Ohnuma T, Minamide Y, Mizuno A, Nishiyama T, Nagashima S, Kanamaru M, Hiratsuka A, Watabe T, Uematsu T (2005) Dihydropyrimidine dehydrogenase activity in 150 healthy Japanese volunteers and identification of novel mutations. Clin Cancer Res 11:5104–5111

Ridge SA, Sludden J, Brown O, Robertson L, Wei X, Sapone A, Fernandez-Salguero PM, Gonzalez FJ, Vreken P, van Kuilenburg AB, van Gennip AH, McLeod HL (1998a) Dihydropyrimidine dehydrogenase pharmacogenetics in Caucasian subjects. Br J Clin Pharmacol 46:151–156

Ridge SA, Sludden J, Wei X, Sapone A, Brown O, Hardy S, Canney P, Fernandez-Salguero P, Gonzalez FJ, Cassidy J, McLeod HL (1998b) Dihydropyrimidine dehydrogenase pharmacogenetics in patients with colorectal cancer. Br J Cancer 77:497–500

Seck K, Riemer S, Kates R, Ullrich T, Lutz V, Harbeck N, Schmitt M, Kiechle M, Diasio R, Gross E (2005) Analysis of the DPYD gene implicated in 5-fluorouracil catabolism in a cohort of Caucasian individuals. Clin Cancer Res 11:5886–5892

Shestopal SA, Johnson MR, Diasio RB (2000) Molecular cloning and characterization of the human dihydropyrimidine dehydrogenase promoter. Biochim Biophys Acta 1494:162–169

van Kuilenburg AB (2004) Dihydropyrimidine dehydrogenase and the efficacy and toxicity of 5-fluorouracil. Eur J Cancer 40:939–950

van Kuilenburg AB, Haasjes J, Richel DJ, Zoetekouw L, Van Lenthe H, De Abreu RA, Maring JG, Vreken P, van Gennip AH (2000) Clinical implications of dihydropyrimidine dehydrogenase (DPD) deficiency in patients with severe 5-fluorouracil-associated toxicity: identification of new mutations in the DPD gene. Clin Cancer Res 6:4705–4712

van Kuilenburg AB, Dobritzsch D, Meinsma R, Haasjes J, Waterham HR, Nowaczyk MJ, Maropoulos GD, Hein G, Kalhoff H, Kirk JM, Baaske H, Aukett A, Duley JA, Ward KP, Lindqvist Y, van Gennip AH (2002) Novel disease-causing mutations in the dihydropyrimidine dehydrogenase gene interpreted by analysis of the three-dimensional protein structure. Biochem J 364:157–163

Vreken P, Van Kuilenburg AB, Meinsma R, van Gennip AH (1997) Dihydropyrimidine dehydrogenase (DPD) deficiency: identification and expression of missense mutations C29R, R886H and R235W. Hum Genet 101:333–338

Wei X, Elizondo G, Sapone A, McLeod HL, Raunio H, Fernandez-Salguero P, Gonzalez FJ (1998) Characterization of the human dihydropyrimidine dehydrogenase gene. Genomics 51:391–400

Yamaguchi K, Arai Y, Kanda Y, Akagi K (2001) Germline mutation of dihydropyrimidine dehydrogenese gene among a Japanese population in relation to toxicity to 5-Fluorouracil. Jpn J Cancer Res 92:337–342

Zhang K, Qin Z, Chen T, Liu JS, Waterman MS, Sun F (2005) HapBlock: haplotype block partitioning and tag SNP selection software using a set of dynamic programming algorithms. Bioinformatics 21:131–134

Zhu AX, Puchalski TA, Stanton VP Jr, Ryan DP, Clark JW, Nesbitt S, Charlat O, Kelly P, Kreconus E, Chabner BA, Supko JG (2004) Dihydropyrimidine dehydrogenase and thymidylate synthase polymorphisms and their association with 5-fluorouracil/leucovorin chemotherapy in colorectal cancer. Clin Colorectal Cancer 3:225–234

Int. J. Cancer: **122**, 2148–2153 (2008)

Synergistic antitumor activity of the novel SN-38-incorporating polymeric micelles, NK012, combined with 5-fluorouracil in a mouse model of colorectal cancer, as compared with that of irinotecan plus 5-fluorouracil

Takako Eguchi Nakajima[1,2], Masahiro Yasunaga[2], Yasuhiko Kano[3], Fumiaki Koizumi[4], Ken Kato[1], Tetsuya Hamaguchi[1], Yasuhide Yamada[1], Kuniaki Shirao[1], Yasuhiro Shimada[1] and Yasuhiro Matsumura[2]*

[1]*Gastrointestinal Oncology Division, National Cancer Center Hospital, Tokyo, Japan*
[2]*Investigative Treatment Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwa, Chiba, Japan*
[3]*Hematology Oncology, Tochigi Cancer Center, Tochigi, Japan*
[4]*Shien Lab Medical Oncology Division, National Cancer Center Hospital, Tokyo, Japan*

The authors reported in a previous study that NK012, a 7-ethyl-10-hydroxy-camptothecin (SN-38)-releasing nano-system, exhibited high antitumor activity against human colorectal cancer xenografts. This study was conducted to investigate the advantages of NK012 over irinotecan hydrochloride (CPT-11) administered in combination with 5-fluorouracil (5FU). The cytotoxic effects of NK012 or SN-38 (an active metabolite of CPT-11) administered in combination with 5FU was evaluated *in vitro* in the human colorectal cancer cell line HT-29 by the combination index method. The effects of the same drug combinations was also evaluated *in vivo* using mice bearing HT-29 and HCT-116 cells. All the drugs were administered i.v. 3 times a week; NK012 (10 mg/kg) or CPT11 (50 mg/kg) was given 24 hr before 5FU (50 mg/kg). Cell cycle analysis in the HT-29 tumors administered NK012 or CPT-11 *in vivo* was performed by flow cytometry. NK012 exerted more synergistic activity with 5FU compared to SN-38. The therapeutic effect of NK012/5FU was significantly superior to that of CPT-11/5FU against HT-29 tumors ($p = 0.0004$), whereas no significant difference in the antitumor effect against HCT-116 tumors was observed between the 2-drug combinations ($p = 0.2230$). Cell-cycle analysis showed that both NK012 and CPT-11 tend to cause accumulation of cells in the S phase, although this effect was more pronounced and maintained for a more prolonged period with NK012 than with CPT-11. Optimal therapeutic synergy was observed between NK012 and 5FU, therefore, this regimen is considered to hold promise of clinical benefit, especially for patients with colorectal cancer.

Key words: NK012; SN-38; 5-fluorouracil; drug delivery system; colorectal cancer

The 5-year survival rates of colorectal cancer (CRC) have improved remarkably over the last 10 years, accounted for in large part by the extensively investigated agents after 5-fluorouracil (5FU). Irinoetcan hydrochloride (CPT-11), a water-soluble, semi-synthetic derivative of camptothecin, is one such agent that has been shown to be highly effective, and currently represents a key-drug in first- and second-line treatment regimens for CRC. CPT-11 monotherapy, however, has not been shown to yield superior efficacy, including in terms of the median survival time, to bolus 5FU/leucovorin (LV) alone.[1] In 2 Phase III trials, the addition of CPT-11 to bolus or infusional 5FU/LV regimens clearly yielded greater efficacy than administration of 5FU/LV alone, with a doubling of the tumor response rate and prolongation of the median survival time by 2–3 months.[1,2]

CPT-11 is converted to 7-ethyl-10-hydroxy-camptothecin (SN-38), a biologically active and water-insoluble metabolite of CPT-11, by carboxylesterases in the liver and the tumor. SN-38 has been demonstrated to exhibit up to a 1,000-fold more potent cytotoxic activity than CPT-11 against various cancer cells *in vitro*.[3] The metabolic conversion rate is, however, very low, with only <10% of the original volume of CPT-11 being metabolized to SN-38[4,5]; conversion of CPT-11 to SN-38 also depends on genetic interindividual variability of the activity of carboxylesterases.[6] Direct use of SN-38 itself for clinical cancer treatment must be shown to be identical in terms of both efficacy and toxicity.

Some drugs incorporated in drug delivery systems (DDS), such as Abraxane and Doxil, are already in clinical use.[7,8] The clinical benefits of DDS are based on their EPR effect.[9] The EPR effect is based on the pathophysiological characteristics of solid tumor tissues: hypervascularity, incomplete vascular architecture, secretion of vascular permeability factors stimulating extravasation within cancer tissue, and absence of effective lymphatic drainage from the tumors that impedes the efficient clearance of macromolecules accumulated in solid tumor tissues. Several types of DDS can be used for incorporation of a drug. A liposome-based formulation of SN-38 (LE-SN38) has been developed, and a clinical trial to assess its efficacy is now under way.[10,11]

Recently, we demonstrated that NK012, novel SN-38-incorporating polymeric micelles, exerted superior antitumor activity and less toxicity than CPT-11.[12] NK012 is characterized by a smaller size of the particles than LE-SN38; the mean particle diameter of NK012 is 20 nm. NK012 can release SN-38 under neutral conditions even in the absence of a hydrolytic enzyme, because the bond between SN-38 and the block copolymer is a phenol ester bond, which is stable under acidic conditions and labile under mild alkaline conditions. The release rate of SN-38 from NK012 under physiological conditions is quite high; more than 70% of SN-38 is released within 48 hr. We speculated that the use of NK012, in place of CPT-11, in combination with 5FU may yield superior results in the treatment of CRC. In the present study, we evaluated the antitumor activity of NK012 administered in combination with 5FU as compared to that of CPT-11 administered in combination with 5FU against CRC in an experimental model.

Material and methods

Cells and animals

The human colorectal cancer cell lines used, namely, HT-29 and HCT-116, were purchased from the American Type Culture Collection (Rockville, MD). The HT-29 cells and HCT-116 cells were maintained in RPMI 1640 supplemented with 10% fetal bovine serum (Cell Culture Technologies, Gaggenau-Hoerden, Germany), penicillin, streptomycin, and amphotericin B (100 units/mL, 100 μg/mL, and 25 μg/mL, respectively; Sigma, St. Louis, MO) in a humidified atmosphere containing 5% CO_2 at 37°C.

BALB/c *nu/nu* mice were purchased from SLC Japan (Shizuoka, Japan). Six-week-old mice were subcutaneously (s.c.)

***Correspondence to:** Investigative Treatment Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan. Fax: +81-4-7134-6866. E-mail: yhmatsum@east.ncc.go.jp
Received 2 September 2007; Accepted after revision 20 November 2007
DOI 10.1002/ijc.23381
Published online 14 January 2008 in Wiley InterScience (www.interscience.wiley.com).

inoculated with 1×10^6 cells of HT-29 or HCT-116 cell line in the flank region. The length (a) and width (b) of the tumor masses were measured twice a week, and the tumor volume (TV) was calculated as follows: $TV = (a \times b^2)/2$. All animal procedures were performed in compliance with the Guidelines for the Care and Use of Experimental Animals established by the Committee for Animal Experimentation of the National Cancer Center; these guidelines meet the ethical standards required by law and also comply with the guidelines for the use of experimental animals in Japan.

Drugs

The SN-38-incorporating polymeric micelles, NK012, and SN-38 were prepared by Nippon Kayaku (Tokyo, Japan).[12] CPT-11 was purchased from Yakult Honsha (Tokyo, Japan). 5FU was purchased from Kyowa Hakko (Tokyo, Japan).

Cell growth inhibition assay

HT-29 cells were seeded in 96-well plates at a density of 2,000 cells/well in a final volume of 90 μL. Twenty-four hours after seeding, a graded concentration of NK012 or SN-38 was added concurrently with 5FU to the culture medium of the HT-29 cells in a final volume of 100 μL for drug interaction studies. The culture was maintained in the CO_2 incubator for an additional 72 hr. Then, cell growth inhibition was measured by the tetrazolium salt-based proliferation assay (WST assay; Wako Chemicals, Osaka, Japan). WST-1 labeling solution (10 μL) was added to each well and the plates were incubated at 37°C for 3 hr. The absorbance of the formazan product formed was detected at 450 nm in a 96-well spectrophotometric plate reader. Cell viability was measured and compared to that of the control cells. Each experiment was carried out in triplicate and was repeated at least 3 times. Data were averaged and normalized against the nontreated controls to generate dose-response curves.

Drug interaction analysis

The nature of interaction between NK012 or SN-38 and 5FU against HT-29 cells was evaluated by median-effect plot analyses and the combination index (CI) method of Chou and Talalay.[13] Data analysis was performed using the Calcusyn software (Biosoft, NY, USA). NK012 or SN-38 was combined with 5FU at a fixed ratio that spanned the individual IC_{50} values of each drug. The IC_{50} values were determined on the basis of the dose-response curves using the WST assay. For any given drug combination, the CI is known to represent the degree of synergy, additivity or antagonism. It is expressed in terms of fraction-affected (Fa) values, which represents the percentage of cells killed or inhibited by the drug. Isobologram equations and Fa/CI plots were constructed by computer analysis of the data generated from the median effect analysis. Each experiment was performed in triplicate with 6 gradations and was repeated at least 3 times. The resultant dose-response curves were averaged, to create a single composite dose-response curve for each combination.

In vivo analysis of the effects of NK012 combined with 5FU as compared to those of CPT-11 combined with 5FU

When the mean tumor volumes reached ~93 mm^3, the mice were randomly divided into test groups consisting of 5 mice per group (Day 0). The drugs were administered i.v. *via* the tail vein of the mice. In the groups administered NK012 or 5FU as single agents, the drug was administered on Days 0, 7 and 14. In the combined treatment groups, NK012 or CPT-11 was administered 24 hr before 5FU on Days 0, 7 and 14, according to the previously reported combination schedule for CPT-11 and 5FU.[14] Complete response (CR) was defined as tumor not detectable by palpation at 90 days after the start of treatment, at which time-point the mice were sacrificed. Tumor volume and body weight were measured twice a week. As a general rule, animals in which the tumor volume exceeded 2,000 mm^3 were also sacrificed.

Experiment 1. Evaluation of the effects of NK012 combined with 5FU and determination of the maximum tolerated dose (MTD) of NK012/5FU. By comparing the data between NK012 administered as a single agent and NK012/5FU, we evaluated the effects of the combined regimen against the s.c HT-29 tumors. A preliminary experiment showed that combined administration of NK012 15 mg/kg + 5FU 50 mg/kg every 6 days caused drug-related lethality (data not shown). To determine the MTD, therefore, we set the dosing schedule of the combined regimen at 5 or 10 mg/kg of NK012 + 50 mg/kg of 5FU three times a week.

Experiment 2. Comparison of the antitumor effect of NK012/5FU and CPT-11/5FU. Based on a comparison of the data between NK012/5FU and CPT-11/5FU against the s.c. HT-29 and HCT-116 tumors, we investigated the feasibility of the clinical application of NK012/5FU for the treatment of CRC. CPT-11/5FU was administered three times a week at the respective MTDs of the 2 drugs as previously reported, that is, CPT11 at 50 mg/kg and 5FU at 50 mg/kg, respectively.[14] NK012/5FU was administered once three times a week at the respective MTDs of the 2 drugs determined from Experiment 1.

Cell cycle analysis

Samples from the HT-29 tumors that had grown to 80–100 mm^3 were removed from the mice at 6, 24, 48, 72 and 96 hr after the administration of NK012 alone at 10 mg/kg or CPT-11 alone at 50 mg/kg. The samples were excised, minced in PBS and fixed in 70% ethanol at −20°C for 48 hr. They were then digested with 0.04% pepsin (Sigma chemical Co., St Lous, MO) in 0.1 N HCL for 60 min at 37°C in a shaking bath to prepare single-nuclei suspensions. The nuclei were then centrifuged, washed twice with PBS and stained with 40 μg/mL of propidium iodide (Molecular Probes, OR) in the presence of 100 μg/mL RNase in 1 mL PBS for 30 min at 37°C. The stained nuclei were analyzed with B-D FACSCalibur (BD Biosciences, San Jose, CA), and the cell cycle distribution was analyzed using the Modfit program (Verity Software House Topsham, ME).

Statistical analyses

Data were expressed as mean ± SD. Data were analysed with Student's t test when the groups showed equal variances (F test), or Welch's test when they showed unequal variances (F test). $p < 0.05$ was regarded as statistically significant. All statistical tests were 2-sided.

Results

Antiproliferative effects of NK012 or SN-38 administered in combination with 5FU

Figure 1*a* shows the dose-response curves for NK012 alone, 5FU alone and a combination of the two. The IC_{50} levels of NK012 and 5FU against the HT-29 cells were 39 nM and 1 μM, respectively, and the IC_{50} level of SN-38 was 14 nM (data not shown). Based on these data, the molar ratio of NK012 or SN-38:5FU of 1:1,000 was used for the drug combination studies.

Figures 1*b* and 1*c* show the median-effect and the combination index plots. Combination indices (CIs) of <1.0 are indicative of synergistic interactions between 2 agents; additive interactions are indicated by CIs of 1.0, and antagonism by CIs of >1.0. Figure 1*c* shows the combination index for NK012 and 5FU, when 2 drugs are supposed to be mutually exclusive. Marked synergism was observed between Fa 0.2 and 0.6. Theoretically, the CI method is the most reliable around an Fa of 0.5, suggesting synergistic effects of the combination of NK012 and 5FU. This synergistic effect was more evident than that of SN-38/5FU (Fig. 1*d*).

In vivo *effect of combined NK012 and 5FU*

Experiment 1. Dose optimization and effect of combined NK012 and 5FU against HT-29 tumors. Comparison of the relative tumor volumes on Day 40 revealed significant differences between

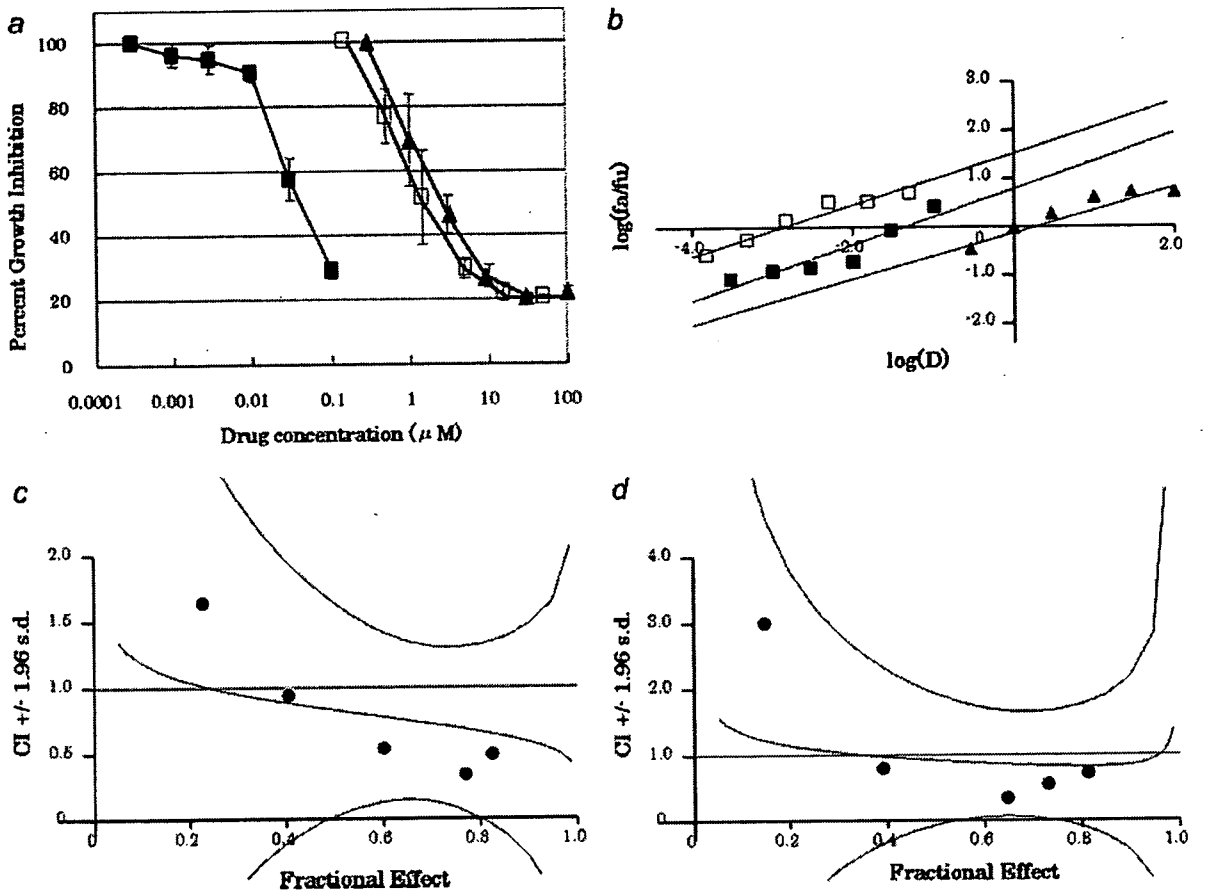


FIGURE 1 – Interaction of NK012 and 5FU *in vitro*. (*a*) Dose-response curves for NK012 alone (■), 5FU alone (▲) and their combination (□) against HT-29 cells. HT-29 cells were seeded at 2,000 cells/well. Twenty-four hours after seeding, a graded concentration of NK012 or 5FU was added to the culture medium of the HT-29 cells. Cell growth inhibition was measured by WST assay after 72 hr of treatment. Cell viability was measured and compared with that of the control cells. Each experiment was carried out independently and repeated at least 3 times. Points, mean of tripricates; bars, SD. (*b*) Median effect plot for the interaction of NK012 and 5FU. (*c*, *d*) Combination index for the interaction as a function of the level of effect (fraction effect = 0.5 is the IC_{50}). The straight line across the CI value of 1.0 indicates additive effect and CIs above and below indicate antagonism and synergism, respectively. The molar ratio of NK012/5FU (*c*) or SN-38/5FU (*d*) at 1:1,000 was tested by CI analysis. Black circles represent the CIs of the actual data points, solid lines represent the computer-derived CIs at effect levels ranging from 10 to 100% inhibition of cell growth, and the dotted lines represent the 95% confidence intervals.

those in the mice administered NK012 alone and those administered NK012/5FU at 5 mg/kg of NK012 ($p = 0.018$) (Fig. 2*a*). Although there was no statistically significant difference in the relative tumor volume measured on Day 54 between the mice administered NK012 alone and NK012/5FU at 10 mg/kg of NK012 ($p = 0.3050$), a trend of superior antitumor effect was demonstrated in the group treated with NK012/5FU at 10 mg/kg of NK012 (Fig. 2*a*). The CR rates were 20, 40 and 60% for 5 mg/kg NK012 + 50 mg/kg 5FU, 10 mg/kg NK012 alone and 10 mg/kg NK012 + 50 mg/kg 5FU, respectively. The schedule of 10 mg/kg NK012 + 50 mg/kg 5FU resulted in no remarkable toxicity in terms of body weight changes, and these doses were determined as representing the MTDs (Fig. 2*b*).

Experiment 2. Comparison of the antitumor effect of combined NK012/5FU and CPT-11/5FU against HT-29 and HCT-116 tumors. The therapeutic effect of NK012/5FU on Day 60 was significantly superior to that of CPT-11/5FU against the HT-29 tumors ($p = 0.0004$) (Fig. 3*a*). A more potent antitumor effect, namely, a 100% CR rate, was obtained in the NK012/5FU group as compared to the 0% CR rate in the CPT-11/5FU group. Although no statistically significant difference in the relative tumor volume on Day 61 was demonstrated between the NK012/5FU and CPT-11/5FU in the case of the HCT-116 tumors ($p = 0.2230$), a trend of superior antitumor effect against these tumors was observed in the NK012/5FU treatment group (Fig. 3*b*). The CR rates for the case of the HCT-116 tumors were 0% in both NK012/5FU and CPT-11/5FU groups.

Specificity of cell cycle perturbation

We studied the differences in the effects between NK012 10 mg/kg and CPT-11 50 mg/kg on the cell cycle (Fig. 4*a*). The data indicated that both NK012 and CPT-11 tended to cause accumulation of cells in the S phase, although the effect of NK012 was stronger and maintained for a more prolonged period than that of CPT-11; the maximal percentage of S-phase cells in the total cell population in the tumors was 34% at 24 hr after the administration of CPT-11, whereas it was 39% at 48 hr after the administration of NK012 (Figs. 4*b*, and 4*c*).

Discussion

Our primary endpoint was to clarify the advantages of NK012 over CPT-11 administered in combination with 5FU. We demonstrated that combined NK012 and 5FU chemotherapy exerts more

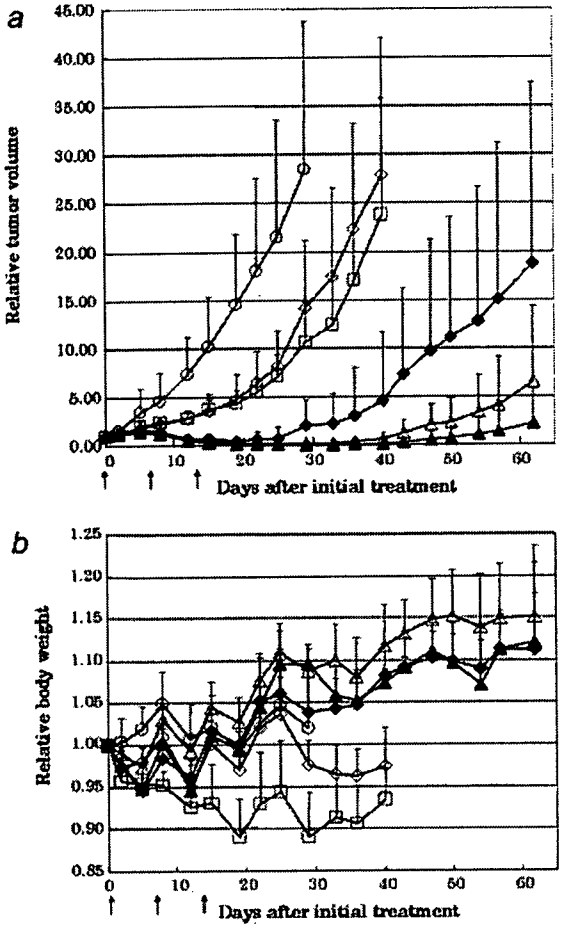


FIGURE 2 – Effect of NK012 alone or NK012 in combination with 5FU against HT-29 tumor-bearing mice. Points, mean; bars, SD. (*a*) Antitumor effect of each regimen on Days 0, 7 and 14. (○) control, (□) 5FU 50 mg/kg alone, (◇) NK012 5 mg/kg alone, (◆) NK012 5 mg/kg 24 hr before 5FU 50 mg/kg, (△) NK012 10 mg/kg alone, (▲) NK012 10 mg/kg 24 hr before 5FU 50 mg/kg. (*b*) Changes in the relative body weight. Data were derived from the same mice as those used in the present study.

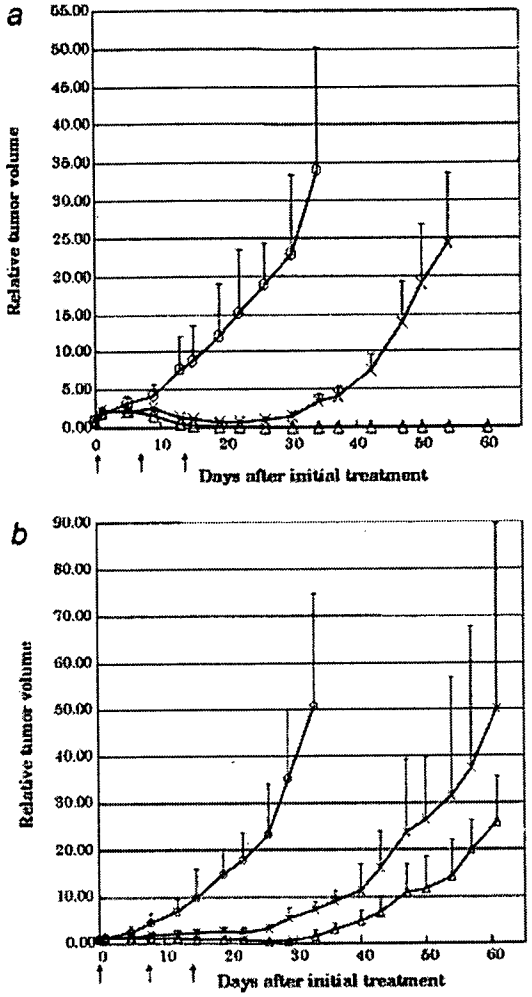


FIGURE 3 – Effect of NK012/5FU as compared with that of CPT11/5FU against HT-29 (*a*) or HCT-116 (*b*) tumor-bearing mice. Antitumor effect of each schedule on Days 0, 7 and 14. (○) control, (×) CPT-11 50 mg/kg 24 hr before 5FU 50 mg/kg, (△) NK012 10 mg/kg 24 hr before 5FU 50 mg/kg. Points, mean; bars, SD.

synergistic activity *in vitro* and significantly greater antitumor activity against human CRC xenografts as compared to CPT-11/5FU. The combination of NK012 and 5FU is considered to hold promise of clinical benefit for patients with CRC.

CPT-11, a topoisomerase-I inhibitor, and 5FU, a thymidilate synthase inhibitor, have been demonstrated to be effective agents for the treatment of CRC. A combination of these 2 drugs has also been demonstrated to be clearly more effective than either CPT-11 or 5FU/LV administered alone *in vivo* and in clinical settings.[1,2,14] Administration of 5FU by infusion with CPT-11 was shown to be associated with reduced toxicity and an apparent improvement in survival as compared to that of administration of the drug by bolus injection with CPT-11.[1,2] This synergistic enhancement may result from the mechanism of action of the 2 drugs; CPT-11 has been reported to cause accumulation of cells in the S phase, and 5FU infusion is known to cause DNA damage specifically in cells of the S phase.[14] On the basis of this background, our results suggesting the more pronounced and more prolonged accumulation of the tumor cells in the S phase caused by NK012 as compared with that by CPT-11 may explain the more effective synergy of the former administered with 5FU infusion. This may be attributable to accumulation of NK012 due to the enhanced permeability and retention (EPR) effect.[9] It is also speculated that NK012 allows sustained release of free SN-38, which may move more freely in the tumor interstitium.[15] Otherwise NK012 itself could internalize into cells to localize in several cytoplasmic organelles as reported by Savic *et al.*[16] These characteristics of NK012 may be responsible for its more potent antitumor activity observed in this study, because CPT-11 has been reported to show time-dependent growth-inhibitory activity against the tumor cells.[17]

The major dose-limiting toxicities of CPT-11 are diarrhea and neutropenia. SN-38, the active metabolite of CPT-11, may cause CPT-11-related diarrhea as a result of mitotic -inhibitory activity.[18] Because it undergoes significant biliary excretion, SN-38 may have a potentially long residence time in the gastrointestinal tract that may be associated with prolonged diarrhea.[19,20] In our previous report, we evaluated the tissue distribution of SN-38 after administration of an equimolar amount of NK012 (20 mg/kg) and CPT-11 (30 mg/kg), and found no difference in the level of SN-38 accumulation in the small intestine.[12] A significant antitumor effect of NK012 with a lower incidence of diarrhea was also dem-

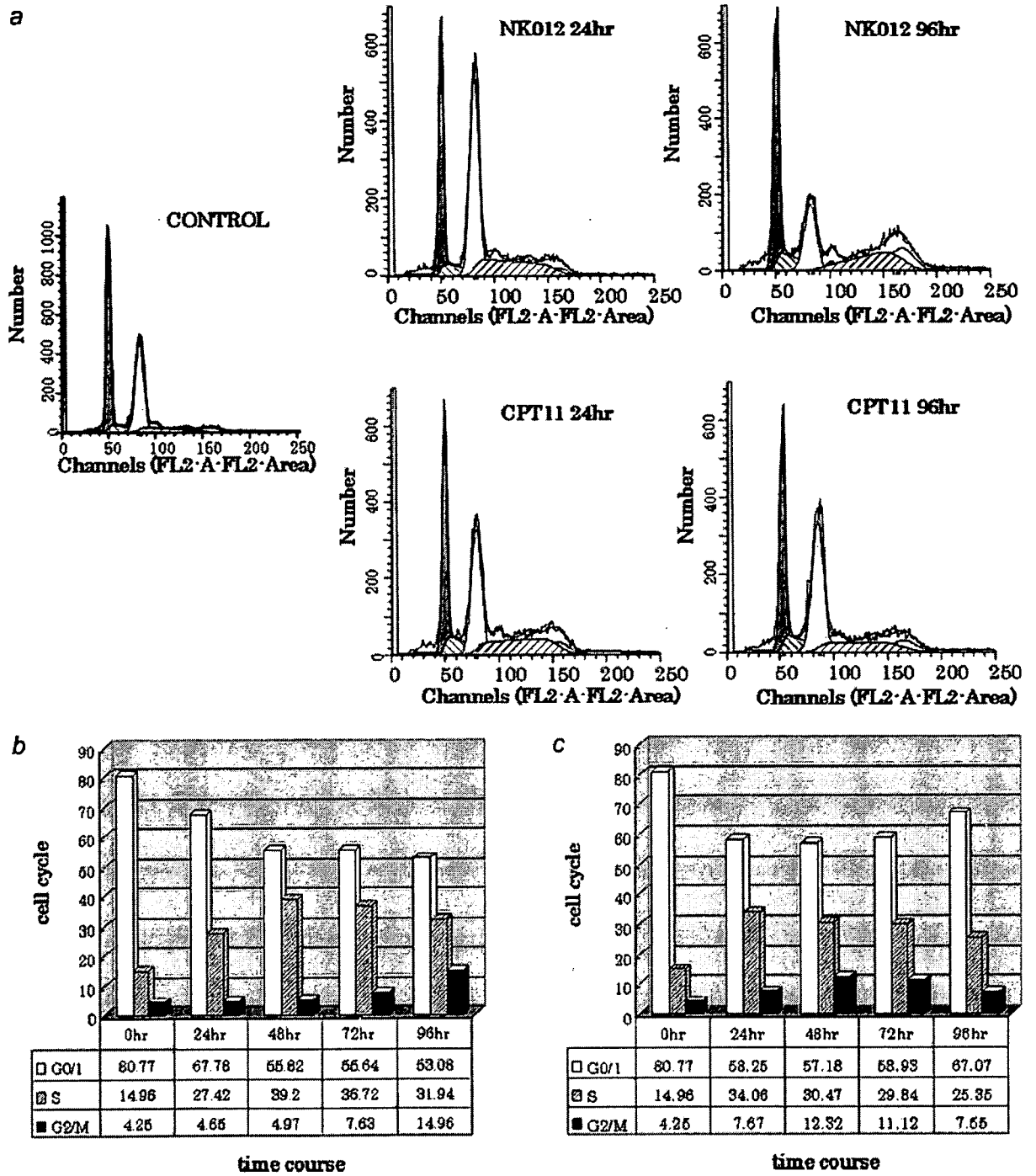


FIGURE 4 – Cell cycle analysis of HT-29 tumor cells collected 24, 48, 72 and 96 hr after administration of NK012 at 10 mg/kg alone or CPT-11 at 50 mg/kg alone using the Modfit program (Verity Software House Topsham, ME). (*a*) Cell cycle analysis of HT-29 tumor cells 24 and 96 hr after administration of NK012 at 10 mg/kg or CPT-11 at 50 mg/kg, respectively. (*b*) Cell cycle distribution of tumor cells 0, 24, 48, 72 and 96 hr after treatment with NK012 at 10 mg/kg. (*c*) Cell cycle distribution of tumor cells 0, 24, 48, 72 and 96 hr after treatment with CPT-11 at 50 mg/kg.

onstrated as compared to that observed with CPT-11 in a rat mammary tumor model.[21] Combined administration of CPT-11 with 5FU/LV infusion appears to be associated with acceptable toxicity in patients with CRC. In addition, no significant difference in the frequency of Grade 3/4 diarrhea was noted between patients treated with FOLFIRI (CPT-11 regimen with bolus and infusional 5FU/LV) and those treated with FOLFOX6 (oxaliplatin regimen with bolus and infusional 5FU/LV).[22,23] Our *in vivo* data actually revealed no severe body weight loss in the NK012/5FU group. Consequently, we expect that the NK012/5FU regimen, especially

with infusional 5FU, may be an attractive arm for a Phase III trial in CRC, with CPT-11/5FU as the control arm. We have already initiated a Phase I trial of NK012 in patients with advanced solid tumors based on the data suggesting higher efficacy and lower toxicity of this preparation than CPT-11 *in vivo*.[12]

In conclusion, we demonstrated that combined NK012 and 5FU chemotherapy exerts significantly greater antitumor activity against human CRC xenografts as compared to CPT-11/5FU, indicating the necessity of clinical evaluation of this combined regimen.

References

1. Saltz LB, Douillard JY, Pirotta N, Alakl M, Gruia G, Awad L, Elfring GL, Locker PK, Miller LL. Irinotecan plus fluorouracil/leucovorin for metastatic colorectal cancer: a new survival standard. Oncologist 2001;6:81–91.
2. Douillard JY, Cunningham D, Roth AD, Navarro M, James RD, Karasek P, Jandik P, Iveson T, Carmichael J, Alakl M, Gruia G, Awad L, et al. Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. Lancet 2000;355:1041–7.
3. Takimoto CH, Arbuck SG. Topoisomerase I targeting agents: the camptothecins. In: Chabner BA, Lango DL, eds. Cancer chemotherapy and biotherapy: principal and practice, 3rd ed. Philadelphia, PA: Lippincott Williams and Wilkins, 2001. 579–646.
4. Slatter JG, Schaaf LJ, Sams JP, Feenstra KL, Johnson MG, Bombardt PA, Cathcart KS, Verburg MT, Pearson LK, Compton LD, Miller LL, Baker DS, et al. Pharmacokinetics, metabolism, and excretion of irinotecan (CPT-11) following I.V. infusion of [(14)C]CPT-11 in cancer patients. Drug Metab Dispos 2000;28:423–33.
5. Rothenberg ML, Kuhn JG, Burris HA, III, Nelson J, Eckardt JR, Tristan-Morales M, Hilsenbeck SG, Weiss GR, Smith LS, Rodriguez GI, Rock MK, Von Hoff DD. Phase I and pharmacokinetic trial of weekly CPT-11. J Clin Oncol 1993;11:2194–204.
6. Guichard S, Terret C, Hennebelle I, Lochon I, Chevreau P, Fretigny E, Selves J, Chatelut E, Bugat R, Canal P. CPT-11 converting carboxylesterase and topoisomerase activities in tumour and normal colon and liver tissues. Br J Cancer 1999;80:364–70.
7. Gradishar WJ, Tjulandin S, Davidson N, Shaw H, Desai N, Bhar P, Hawkins M, O'Shaughnessy J. Phase III trial of nanoparticle albumin-bound paclitaxel compared with polyethylated castor oil-based paclitaxel in women with breast cancer. J Clin Oncol 2005;23:7794–803.
8. Muggia FM. Liposomal encapsulated anthracyclines: new therapeutic horizons. Curr Oncol Rep 2001;3:156–62.
9. Matsumura Y, Maeda H. A new concept for macromolecular therapeutics in cancer chemotherapy: mechanism of tumoritropic accumulation of proteins and the antitumor agent smancs. Cancer Res 1986;46:6387–92.
10. Zhang JA, Xuan T, Parmar M, Ma L, Ugwu S, Ali S, Ahmad I. Development and characterization of a novel liposome-based formulation of SN-38. Int J Pharm 2004;270:93–107.
11. Kraut EH, Fishman MN, LoRusso PM, Gorden MS, Rubin EH, Haas A, Fetterly GJ, Cullinan P, Dul JL, Steinberg JL. Final result of a phase I study of liposome encapsulated SN-38 (LE-SN38): safety, pharmacogenomics, pharmacokinetics, and tumor response [abstract 2017]. Proc Am Soc Clin Oncol 2005;23:139S.
12. Koizumi F, Kitagawa M, Negishi T, Onda T, Matsumoto S, Hamaguchi T, Matsumura Y. Novel SN-38-incorporating polymeric micelles, NK012, eradicate vascular endothelial growth factor-secreting bulky tumors. Cancer Res 2006;66:10048–56.
13. Chou TC, Talalay P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. Adv Enzyme Regul 1984;22:27–55.
14. Azrak RG, Cao S, Slocum HK, Toth K, Durrani FA, Yin MB, Pendyala L, Zhang W, McLeod HL, Rustum YM. Therapeutic synergy between irinotecan and 5-fluorouracil against human tumor xenografts. Clin Cancer Res 2004;10:1121–9.
15. Jain RK. Barriers to drug delivery in solid tumors. Sci Am 1994; 271:58–65.
16. Savic R, Luo L, Eisenberg A, Maysinger D. Micellar nanocontainers distribute to defined cytoplasmic organelles. Science 2003;300:615–18.
17. Kawato Y, Aonuma M, Hirota Y, Kuga H, Sato K. Intracellular roles of SN-38, a metabolite of the camptothecin derivative CPT-11, in the antitumor effect of CPT-11. Cancer Res 1991;51:4187–91.
18. Slater R, Radstone D, Matthews L, McDaid J, Majeed A. Hepatic resection for colorectal liver metastasis after downstaging with irinotecan improves survival. Proc Am Soc Clin Oncol 2003;22(abstract 1287).
19. Araki E, Ishikawa M, Iigo M, Koide T, Itabashi M, Hoshi A. Relationship between development of diarrhea and the concentration of SN-38, an active metabolite of CPT-11, in the intestine and the blood plasma of athymic mice following intraperitoneal administration of CPT-11. Jpn J Cancer Res 1993;84:697–702.
20. Atsumi R, Suzuki W, Hakusui H. Identification of the metabolites of irinotecan, a new derivative of camptothecin, in rat bile and its biliary excretion. Xenobiotica 1991;21:1159–69.
21. Onda T, Nakamura I, Seno C, Matsumoto S, Kitagawa M, Okamoto K, Nishikawa K, Suzuki M. Superior antitumor activity of NK012, 7-ethyl-10-hydroxycamptoyhecin-incorporating micellar nanoparticle, to irrinotecan. Proc Am Assoc Cancer Res 2006;47:720s(abstract 3062).
22. Tournigand C, Andre T, Achille E, Lledo G, Flesh M, Mery-Mignard D, Quinaux E, Couteau C, Buyse M, Ganem G, Landi B, Colin P, et al. FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. J Clin Oncol 2004;22:229–37.
23. Colucci G, Gebbia V, Paoletti G, Giuliani F, Caruso M, Gebbia N, Carteni G, Agostara B, Pezzella G, Manzione L, Borsellino N, Misino A, et al. Phase III randomized trial of FOLFIRI versus FOLFOX4 in the treatment of advanced colorectal cancer: a multicenter study of the Gruppo Oncologico Dell'Italia Meridionale. J Clin Oncol 2005; 23:4866–75.

Expression profiling of fecal colonocytes for RNA-based screening of colorectal cancer

SATOSHI YAJIMA[1,2], MIE ISHII[3], HISAYUKI MATSUSHITA[4], KAZUHIKO AOYAGI[1], KAZUHIKO YOSHIMATSU[5], HIRONORI KANEKO[2], NOBUKO YAMAMOTO[3], TATSUO TERAMOTO[2], TERUHIKO YOSHIDA[1], YASUHIRO MATSUMURA[4] and HIROKI SASAKI[1]

[1]Genetics Division, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045; [2]Division of General and Gastroenterological Surgery (Omori), Department of Surgery, School of Medicine, Faculty of Medicine, Toho University, Omori Nishi 6-11-1, Ohta-ku, Tokyo 143-8541; [3]Medical Engineering Development Center, Canon Inc., Shimomaruko 3-30-2, Ohta-ku, Tokyo 146-8501; [4]Investigate Treatment Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwanoha 6-5-1, Kashiwa, Chiba 277-8577; [5]Medical Center East, Tokyo Women's Medical University, School of Medicine, Nishiogu 2-1-10, Arakawa-ku, Tokyo 116-8567, Japan

Received May 21, 2007; Accepted July 18, 2007

Abstract. The early detection of colorectal cancer originating from any part of the colorectum is desirable because this cancer can be cured surgically if diagnosed early. We searched for marker genes for a fecal RNA-based colorectal cancer screening method by comparison of genome-wide expression profiles among cancerous and non-cancerous tissues, and healthy volunteer- and cancer patient-derived colonocytes from the feces, and the peripheral blood. Of 14,564 genes, only 3 (PAP, REG1A, and DPEP1) were selectable as final candidates which were expressed frequently at any stage of this cancer and were suppressed in non-cancerous tissues and also in the peripheral blood and colonocytes of healthy volunteers. Next, we directly compared fecal RNA-expression profiles between colorectal cancer patients and healthy volunteers, and found that most of the genes (92%) expressed in the colonocytes of the cancer patients were not expressed in those of the healthy volunteers. Six genes (SEPP1, RPL27A, ATP1B1, EEF1A1, SFN, and RPS11) selected randomly from 85 cancer patient-derived colonocyte-specific genes were evaluated. In total, reverse transcription-polymerase chain reaction or focused microarray of all those 9 genes detected 18 (78%) of 23 curable colorectal cancers (Dukes stages A-C), 9 or 10 (64% or 71%) of 14 early cancers with no lymph node metastasis (Dukes stage A or B) and 4 (80%) of 5 right-sided cancers. Our extensive gene list provides other markers for fecal RNA-based colorectal cancer screening.

Correspondence to: Dr Hiroki Sasaki, Genetics Division, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan
E-mail: hksasaki@gan2.res.ncc.go.jp

Key words: expression profiling, colonocyte, colorectal cancer screening, microarray

Introduction

Colorectal cancer is a common malignancy which is curable by surgical resection if diagnosed at a sufficiently early stage (stage I/Dukes stage A or stage II/Dukes stage B). Five-year survival rates on surgical resection, for example, at Dukes stage A, Dukes stage B and stage III/Dukes stage C are 95%, 80% and 50-60%, respectively. For stage IV/Dukes stage D, curative resection is impossible. Therefore, early detection of this cancer originating from any part of the colorectum is desired. For mass cancer screenings, a simple, economic, and noninvasive method of cancer detection is required. The Hemoccult test is currently used in many countries for this purpose (1-5). However, this test is nonspecific and is not sufficiently sensitive to detect early-stage cancer, although a higher sensitivity has been reported for the advanced stage (6).

For fecal DNA-based colorectal cancer screening, in 1992, Sidransky first reported Ras oncogene mutations in the fecal DNA of patients with curable colorectal cancer (7). To date, many screening methods based on mutated DNA detection in the feces have been reported (8-19). These methods, however, are time-consuming and are not sufficiently sensitive. The major reason for this inaccuracy is the fact that fecal DNAs are derived from an enormous number and variety of bacteria and normal living cells including normal colorectal mucus cells, lymphocytes, red blood cells and anal squamous cells. Immunocytochemical analysis provides a simple method; however, this method is insensitive because only the surface portion of the feces can be assayed. On the other hand, Tarin and colleagues first reported that cancer-specific CD44 splicing variants are useful for fecal RNA-based colorectal cancer screening (20,21). By the use of the repetition of the Percoll centrifugation method for isolating the colonocytes from feces, we have also demonstrated that unusual CD44 variants could be targets for cancer-detection using feces (22). However, the method is found to distort the morphology of

colonocytes and to have a low retrieval rate. Accordingly, the sensitivity of this mRNA-based method also appears to be insensitive.

In any method for colorectal cancer detection using feces, an effective method which allows the simple isolation of the colonocytes from not only the surface but also the central portion of the feces while maintaining the initial morphology is needed. Recently, we successfully developed a new, very effective method that is based on the filtration of the homogenates of feces and magnetic cell sorting (MACS) with an epithelial cell-specific antibody, which we here abbreviated to FMCI (filtration and MACS-based colonocyte isolation) (23). It has been shown that this method can provide a high quality of colonocyte DNA or RNA for molecular biological analysis and also provide the colonocyte with its original morphology for cytology. Considering the advantage in the use of FMCI, we here report expression profiling of colonocytes for fecal RNA-based detection of curable colorectal cancer and a sensitive focused microarray assay that uses multiple marker genes for detecting minimal cancer cells in the feces of patients with colorectal cancer.

Materials and methods

Clinical materials. This study protocol was reviewed and approved by the Institutional Review Board of the National Cancer Center, Tokyo. Written informed consent was obtained from all the patients and healthy volunteers. Before surgical resection, stool samples were obtained from 23 patients with colorectal cancer (Dukes stages A-C), for which curable resection is possible, and from 15 healthy volunteers a few weeks after they had undergone a total colonoscopy. Naturally evacuated feces from subjects who had not taken laxatives were used as stool samples. Each patient was instructed to evacuate into a polystyrene disposable tray (AS one, Osaka, Japan) measuring 5x10 cm in size. Preparation of the stool samples for examination was conducted within 1-6 h after the evacuation. Tissue samples were obtained from the surgically resected specimens of colorectal cancer patients, and were snap frozen in liquid nitrogen and stored at -80°C until use. RNA of the tissues was extracted by using an Isogen kit (Nippon Gene, Toyama, Japan). The peripheral blood samples of 58 healthy volunteers and their RNA were prepared as in our previous report (24).

Isolation of exfoliated cells from feces. The procedure is detailed in our previous report (23). In brief, approximately 5-10 g of feces was used to isolate exfoliated cells. Feces were collected in Stomacher Lab Blender bags (Seward, Thetford, UK). The stool samples were homogenized with a buffer (200 ml) consisting of Hank's solution, 25 mM Hepes (pH 7.35), and 10% fetal bovine serum at 200 rpm for 1 min using a Stomacher (Seward). The homogenates were then filtered through a nylon filter (pore size, 512 μm), followed by division into 5 portions (40 ml each). Subsequently, 40 μl of magnetic beads coated with a mouse IgG1 monoclonal antibody (mAb Ber-EP4) specific for the glycopolypeptide membrane antigen Ep-CAM, which is expressed on most normal and neoplastic human epithelial cells (Dynal, Oslo, Norway), was added to each portion, and the mixtures were incubated for 30 min under gentle rolling in a mixer at room temperature. After 15-min shaking, the colonocytes were recovered from 5 tubes. The colonocytes from a single tube were stored at -80°C for RNA extraction. The colonocyte RNA was extracted by using an Isogen kit (Nippon Gene, Toyama, Japan).

Microarray analysis. We used human U133A Gene Chip (Affymetrix, Santa Clara, CA) for genome-wide expression profiling of mRNAs corresponding to 14,564 genes, 18,445 transcripts including splicing variants, and 22,215 probe sets. The procedures were conducted according to the supplier's protocols. Briefly, 10 μg of fragmented cRNA was hybridized to the microarrays in 200 μl of a hybridization cocktail at 45°C for 16 h in a rotisserie oven set at 60 rpm. The arrays were then washed with a nonstringent wash buffer (6X SSPE) at 25°C, followed by a stringent wash buffer [100 mM MES (pH 6.7), 0.1 M NaCl, and 0.01% Tween-20] at 50°C, stained with streptavidin phycoerythrin (Invitrogen, Carlsbad, CA), washed again with 6X SSPE, stained with biotinylated anti-streptavidin IgG, followed by a second staining with streptavidin phycoerythrin and a third wash with 6X SSPE. The arrays were scanned using a GeneArray scanner (Affymetrix) at 3-μm resolution, and the expression value (average difference: AD) of each gene was calculated using GeneChip Analysis Suite version 5.0 software (Affymetrix). The mean of AD values in each experiment was 1000 to reliably compare variable multiple arrays.

Reverse transcription-polymerase chain reaction (RT-PCR). RT-PCR on colonocyte RNA was carried out using primer sets designed for detecting the 3' side of cDNA of each gene. Primers were 5'-ACCAGTGTGAGGACTCACCC-3' and 5'-TGCTCTTTAAAGCCTTAGGCC-3' for PAP; 5'-AGCAATTACAACGGAGTCAA-3' and 5'-TCCAAAGACTGGGGTAGGT-3' for REGIA; 5'-TCTCTCCTGTGAAACCTGGG-3' and 5'-AAGGGGTGTTGCTTTTATTGC-3' for DPEP1; 5'-ATTAGCAGTTTAGAATGGAGG-3' and 5'-CTGTATCCAATTCTGTACTGC-3' for SEPP1; 5'-TGGGCTGCCAACATGCCATC-3' and 5'-TGTAGTAGCCCGATCGCACC-3' for RPL27A; 5'-GGCAAGCGAGATGAAGATAAGG-3' and 5'-AGGTCCCATACGTATGACAG-3' for ATP1B1; 5'-AGACTATCCACCTTTGGGTCG-3' and 5'-GATGCATTGTTATCATTAACCAGTC-3' for EEF1A; 5'-TTGAGCGCACCTAACCACTGGT-3' and 5'-GAGAGGAAACATGGTCACACCCA-3' for SFN, and 5'-ACATTCAGACTGAGCGTGCCTA-3' and 5'-GATCTGGACGTCCCTGAAGCA-3' for RPS11. PCR was performed under conditions of 30-35 cycles of 3 steps of temperature, 95°C for 1 min, 55°C for 1 min, and 72°C for 1 min, using the AccuPrime TaqDNA polymerase system (Invitrogen).

Marker gene detection using focused microarray. A focused microarray was constructed by fixing 50-60 mer of oligonucleotide probes on a slide glass using our previously developed Bubble Jet Technology (25). The microarray contained a single spot for each sequence of 9 marker genes (PAP, REG1A, DPEP1, SEPP1, RPL27A, ATP1B1, EEF1A1, SFN, and RPS11) and a control artificial DNA sequence. Each probe sequence used for the microarray is listed in Table I.

Table I. Sequences of primers and probes for focused microarray analysis.

Gene	Forward primers Reverse primers	Probes
PAP	GAGAAGCACAGCATTTCTGAG TGCTCTTTAAAGCCTTAGGCC	TTCCCCCAACCTGACCACCTCATTCTTATCTTTCTTCTGT TTCTTCCTCCCCGCTGTCAT
REG1A	AATCCTGGCTACTGTGTGAG TCCAAAGACTGGGGTAGGT	GACCATCTCTCCAACTCAACTCAACCTGGACACTCTCTT CTCTGCTGAGTTTGCCTTGTT
DPEP1	ACCCATTACGGCTACTCCTC AAGGGGTGTTGCTTTTATTGC	CAGATGCCAGGAGCCCTGCTGCCCACATGCAAGGACCA GCATCTCCTGAGAG
SEPP1	AATTAGCAGTTTAGAATGGAGG CTGTATCCAATTCTGTACTGC	CCATAGTCAATGATGGTTTAATAGGTAAACCAAACCCTA TAAACCTGACCTCCTTTATGG
RPL27A	TGGGCTGCCAACATGCCATC TGTAGTAGCCCGATCGCACC	CCAACTGTCAACCTTGACAAATTGTGGACTTTGGTCAGT GAACAGACACGGGTGAATGCT
ATP1B1	GGCAAGCGAGATGAAGATAAGG AGGTCCCATACGTATGACAG	GAGTGTAAGGCGTACGGTGAGAACATTGGGTACAGTGA GAAAGACCGTTTTCAGGGACGT
EEF1A1	AGACTATCCACCTTTGGGTCG GATGCATTGTTATCATTAACCAGTC	CCACCCCACTCTTAATCAGTGGTGGAAGAACGGTCTCAG AACTGTTTGTTTCAATTGGCC
SFN	TTGAGCGCACCTAACCACTGGT GAGAGGAAACATGGTCACACCCA	CTCTGATCGTAGGAATTGAGGAGTGTCCCGCCTTGTGGC TGAGAACTGGACAGTGG
RPS11	ACATTCAGACTGAGCGTGCCTA GATCTGGACGTCCCTGAAGCA	TCATCCGCCGAGACTATCTGCACTACATCCGCAAGTACA ACCGCTTCGAGAAGCG

Focused microarray analysis consists of 3 steps: i) Cy3-dUTP labeling by multiplex-RT-PCR; ii) hybridization Cy3-labeled cDNA to microarray, and iii) fluorescence scanning (Fig. 3). Using 0.5 to 1 μg of total RNA prepared from colonocytes, reverse transcription was performed with Superscript II (Invitrogen) with T7-oligo dT 24 primer in a total volume of 20 μl according to the manufacturer's protocol. To obtain 5-10 μg of cRNA, T7 transcription was performed. Using 5-10 μg of the cRNA, reverse transcription was performed with Superscript II with random hexamer in a total volume of 20 μl. Multiplex-RT-PCR was performed in two tubes at different PCR cycles: 35 cycles for PAP, REG1A and DPEP1, and 25 cycles for SEPP1, RPL27A, ATP1B1, EEF1A1, SFN, and RPS11. PCR primer sequences are also listed in supplementary Table II. Twenty-five μl of the PCR solution in each tube consisted of 1 μl of template cDNA, primers (6.25 pmol each), 50 μM Cy3-dUTP, 2.5 μl of AccuPrime 10X buffer1 (2 mM dNTP, 15 mM $MgCl_2$) and 1.0 μl of AccuPrime Taq polymerase (Invitrogen). A thermal cycler was set with initial heating at 95°C for 5 min followed by an amplification cycle heated at 95°C for 30 sec, 58°C for 30 sec and 72°C for 40 sec, followed by heating at 72°C for 10 min. The two PCR solutions were mixed and purified with a QIAquick PCR purification kit (Qiagen, Tokyo, Japan). The entire Cy3-labeled cDNA solution (50 μl) was mixed in 120 μl of a hybridization cocktail (6X SSPE containing 900 mM NaCl, 60 mM NaH_2PO_4, and H_2O, and 6 mM EDTA, pH 7.4/ 10% formamide/0.05% SDS) including 0.1 nM Cy3-labeled oligonucleotide which hybridizes the control artificial DNA sequence. By using a hybridization apparatus, HybStation (Genomic Solutions, Ann Arbor, MI), an array was pre-heated to 65°C for 3 min, filled with the hybridization cocktail, and incubated at 92°C for 2 min and then at 55°C for 4 h. Subsequently, the array was washed with 2X SSC, 0.1% SDS at 25°C and then with 2X SSC at 20°C, and rinsed with 0.1X SSC in accordance with a conventional manual, and finally dried by a spin drier. The array was scanned by an apparatus for DNA microarrays, Genepix 4000B (Axon Instruments, Union City, CA) and the fluorescence intensity from each probe spot was obtained after subtracting the background level. This focused microarray assay belongs in a negative or positive assay. However, it is required for determining the cutoff values. In this study, the maximum value of each gene plus 2- or 3-times standard deviation in 7 healthy volunteers was used as the cutoff-value.

Results

Marker gene selection through genome-wide expression profiles of cancer tissues, non-cancerous tissues, and the peripheral blood. In the feces of colorectal cancer patients, living cells other than bacteria include a small amount of cancer cells and normal colorectal mucus cells, lymphocytes, red blood cells and anal squamous cells. It is noted that the content of lymphocytes and red blood cells is increased in the feces of people with hemorrhoids. Therefore, genes that are expressed in almost all cases of early and advanced colorectal cancer and that are not expressed in normal colorectal mucosas,