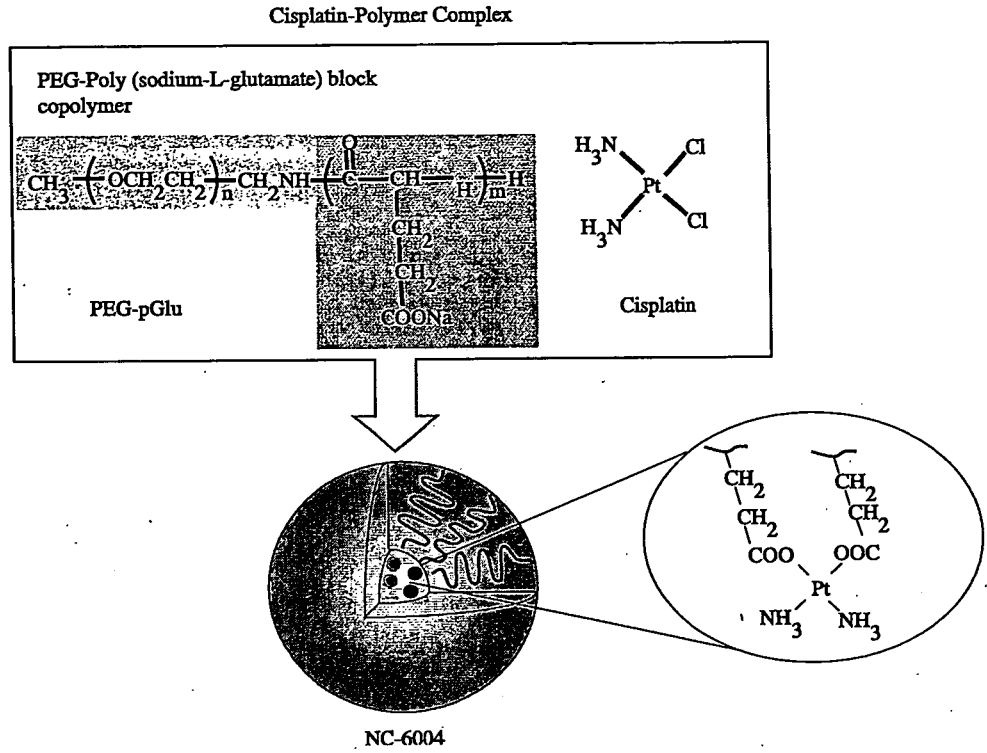

Figure 4. Preparation and characterization of cisplatin-incorporating polymeric micelles (NC-6004). Chemical structures of cisplatin (CDDP) and polyethylene glycol poly(glutamic acid) block copolymers [PEG-P(Glu) block copolymers], and the micellar structures of CDDP-incorporating polymeric micelles (NC-6004).

CDDP, i.e. one-nineteenth and one-seventy fifth, respectively, ($P < 0.01$ and 0.01, respectively).

Regarding the concentration–time profile of platinum in various tissues after i.v. injection of CDDP or NC-6004, all organs measured exhibited the highest concentrations of platinum within 1 h after administration in all animals given CDDP. Furthermore, animals given NC-6004 exhibited the highest tissue concentrations of platinum in the liver and spleen at late time points (24 and 48 h after administration, respectively). However, the concentrations decreased on day 7 after administration. In addition, and in a similar manner to other drugs, which are incorporated in polymeric carriers, NC-6004 demonstrated accumulation in organs of the reticuloendothelial system, e.g. liver and spleen. At 48 h after administration, tissue concentrations of platinum in the liver and spleen were 4.6- and 24.4-fold higher for NC-6004 than for CDDP. On the other hand, a marked increase in tissue platinum concentration was observed immediately after administration in the kidneys of animals given CDDP. Renal platinum concentration at 10 min and 1 h after administration were 11.6- and 3.1-fold lower, respectively, in animals given NC-6004 than in animals given CDDP. Furthermore, the maximum concentration (C_{max}) in the kidney was 3.8-fold lower at the time of NC-6004 administration than at the time of CDDP administration.

Regarding the tumour accumulation of platinum, tumour concentrations of platinum peaked at 10 min after administration of CDDP. On the other hand, tumour concentrations of platinum peaked at 48 h after administration of NC-6004. The maximum concentration (C_{max}) in tumour was 2.5-fold higher for NC-6004 than for CDDP ($P < 0.001$). Furthermore, the tumour AUC was 3.6-fold higher for NC-6004 than for CDDP (81.2 and 22.6 μg/ml h in animals given NC-6004 and CDDP, respectively).

In vivo *antitumour activity*

BALB/c nude mice implanted with a human gastric cancer cell line MKN-45 showed decreased tumour growth rates after i.v. injection of CDDP and NC-6004. In the administration of CDDP, the CDDP 5 mg/kg administration group showed a significant decrease ($P < 0.01$) in tumour growth rate as compared with the control group. However, the NC-6004 administration groups at the same dose levels as CDDP showed no significant difference in tumour growth rate. Regarding time-course changes in body weight change rate, the CDDP 5 mg/kg administration group showed a significant decrease ($P < 0.001$) in body weight as compared with the control group. On the other hand, NC-6004 administration group did not show a decrease in body weight as compared with the control group.

Nephrotoxicity of CDDP and NC-6004

In the CDDP 10 mg/kg administration group, 4 of 12 rats died from toxicity within 7 days after drug administration. No deaths occurred in the NC-6004 10 mg/kg administration group. Regarding renal function, the BUN concentrations on day 7 after the administration of 5% glucose, CDDP 10 mg/kg, and NC-6004 10 mg/kg were 20.8 ± 3.0, 65.3 ± 44.4 and 20 ± 4.5 mg/dl, respectively. The plasma concentrations of creatinine on day 7 after the administration of 5% glucose, CDDP 10 mg/kg, and NC-6004 10 mg/kg were 0.27 ± 0.03, 0.68 ± 0.23 and 0.28 ± 0.04 mg/dl, respectively. The CDDP 10 mg/kg administration group showed significantly higher plasma concentrations of BUN and creatinine as compared with the control group ($P < 0.05$ and 0.001, respectively), with the NC-6004 10 mg/kg administration group ($P < 0.05$ and 0.001, respectively) (Figure 5). Light microscopy indicated tubular dilation with flattening of the lining cells of the tubular epithelium in the kidney from all animals in the CDDP 10 mg/kg administration group. On the other hand, no histopathological change was observed in the kidneys from all animals in the NC-6004 10 mg/kg administration group.

Neurotoxicity of CDDP and NC-6004

Neurophysiological examination revealed that motor nerve conduction velocities (MNCVs) in animals given 5% glucose, CDDP, and NC-6004 were 44.2 ± 3.5, 40.94 ± 5.08 and 40.62 ± 0.63 m/s, respectively. No significant difference was found among the groups with respect to MNCV. Furthermore, sensory nerve conduction velocities (SNCVs) in animals given 5% glucose, CDDP, and NC-6004 were 42.86 ± 8.07, 35.48 ± 4.91 and 43.74 ± 5.3 m/s, respectively. Animals given NC-6004 showed no delay in SNCV as compared with animals given 5% glucose. On the other hand, animals given CDDP showed a significant delay ($P < 0.05$) in SNCV as compared with animals given NC-6004 (Figure 6). The analysis by ICP-MS on sciatic nerve concentrations of platinum could not detect platinum in the sciatic nerve from animals given 5% glucose (data not shown). Sciatic nerve concentrations of platinum

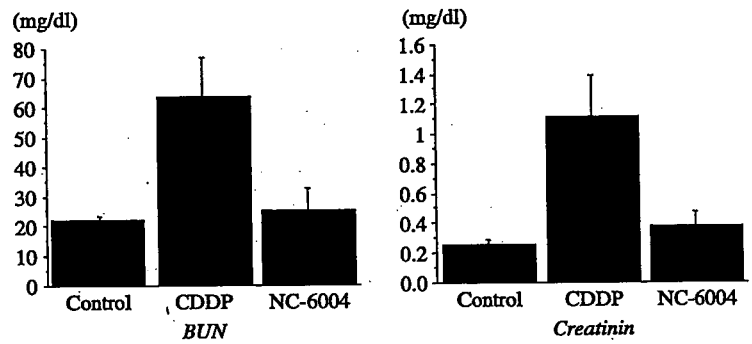

Figure 5. Nephrotoxicity of CDDP and NC-6004. Plasma concentrations of blood urea nitrogen (BUN) and creatinine were measured after a single i.v. injection of 5% glucose ($n = 8$), CDDP at a dose of 10 mg/kg ($n = 12$), NC-6004 at a dose of 10 mg/kg ($n = 13$) on a CDDP basis.

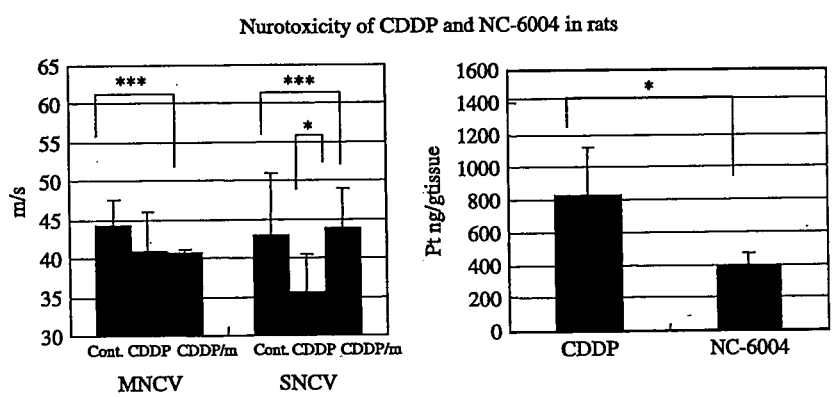

Figure 6. Neurotoxicity of CDDP and NC-6004 in rats. Rats ($n = 5$) were given CDDP (2 mg/kg), NC-6004 (an equivalent dose of 2 mg/kg CDDP), or 5% glucose, all intravenously twice a week, 11 administrations in total. Sensory nerve conduction velocity (SNCV) and motor nerve conduction velocity (MNCV) of the sciatic nerve at week 6 after the initial administration. The platinum concentration in the sciatic nerve. Rats were given CDDP (5 mg/kg, $n = 5$), NC-6004 (an equivalent dose of 5 mg/kg CDDP, $n = 5$), or 5% glucose ($n = 2$), all intravenously twice a week, four administrations in total. On day 3 after the final administration, a segment of the sciatic nerve was removed and the platinum concentration in the sciatic nerve was measured by ICP-MS. The data are expressed as the mean ± SD $*P < 0.05$.

in animals given CDDP and NC-6004 were 827.2 ± 291.3 and 395.5 ± 73.1 ng/g tissue. Therefore, the concentrations were significantly ($P < 0.05$) lower in animals given NC-6004. This finding is believed to be a factor which reduced neurotoxicity following NC-6004 administration as compared with the CDDP administration.

Present situation of a clinical study of NC-6004

A phase 1 clinical trial of NC-6004 is now under way in United Kingdom. Starting dose of NC-6004 was 10 mg/m^2. NC-6004 was administered once every 3 weeks with only 1000 ml water loading. In Japan, a phase 1 trial will be stared soon in the National Cancer Center Hospital.

NK012, SN-38-incorporating micellar nanoparticle

The antitumor plant alkaloid camptothecin (CPT) is a broad-spectrum anticancer agent which targets the DNA topoisomerase I. Although CPT has showed promising antitumor activity *in vitro* and *in vivo* (Gallo et al. 1971; Li et al. 1972), it has not been used clinically because of its low therapeutic efficacy and severe toxicity (Gottlieb et al. 1970; Muggia et al. 1972). Among CPT analogs, irinotecan hydrochloride (CPT-11) has recently been demonstrated to be active against colorectal, lung, and ovarian cancer (Cunningham et al. 1998; Saltz et al. 2000; Noda et al. 2002, Negoro et al. 2003; Bodurka et al. 2003). CPT-11 itself is a prodrug and is converted to 7-ethyl-10-hydroxy-CPT (SN-38), a biologically active metabolite of CPT-11, by carboxylesterases (CEs). SN-38 exhibits up to 1000-fold more potent cytotoxic activity against various cancer cells *in vitro* than CPT-11 (Takimoto and Arbuck 2001). Although CPT-11 is converted to SN-38 in the liver and tumor, the metabolic conversion rate is less than 10% of the original volume of CPT-11 (Rothenberg et al. 1993; Slatter et al. 2000). In addition, the conversion of CPT-11 to SN-38 depends on the genetic inter-individual variability of CE activity (Guichard et al. 1999). Thus, direct use of SN-38 might be of great advantage and attractive for cancer treatment. For the clinical use of SN-38, however, it is essential to develop a soluble form of water-insoluble SN-38. The progress of the manufacturing technology of "micellar nanoparticles" may make it possible to use SN-38 for *in vivo* experiments and further clinical use.

Preparation and characterization of NK012

NK012 is an SN-38-loaded polymeric micelle constructed in an aqueous milieu by the self-assembly of an amphiphilic block copolymers, PEG-PGlu(SN-38). The molecular weight of PEG-PGlu(SN-38) was determined to be approximately 19,000 (PEG segment: 12,000; SN-38-conjugated PGlu segment: 7000). NK012 was obtained as a freeze-dried formulation and contained *ca.* 20% (w/w) of SN-38 (Figure 7). The mean particle size of NK012 is 20 nm in diameter with a relatively narrow range. The releasing rates of SN-38 from NK012 in phosphate buffered saline at 37°C were 57 and 74% at 24 and 48 h, respectively, and that in 5% glucose solution at 37°C were 1 and 3% at 24 and 48 h, respectively. These results indicate that NK012 can release SN-38 under neutral condition even without the presence of a hydrolytic enzyme, and is stable in 5% glucose solution. It is suggested that NK012 is stable before administration and starts to release SN-38, the active component, under physiological conditions after administration.

Cellular sensitivity of NSCLC and colon cancer cells to SN-38, NK012, and CPT-11

The IC_{50} values of NK012 for the cell lines ranged from 0.009 μM (Lovo cells) to 0.16 μM (WiDR cells). The growth-inhibitory effects of NK012 are 43–340 fold more potent than those of CPT-11, whereas the IC_{50} values of NK012 were 2.3–5.8 fold higher than those of SN-38.NK012 exhibited a higher cytotoxic effect against each cell line as compared with CPT-11 (× 43–340 fold sensitivity). On the other hand, the IC_{50} values of NK012 were a little higher than those of SN-38, similar to the cytotoxic feature also reported in a previous study about micellar drugs (Uchino et al. 2005).

Pharmacokinetic analysis of NK012 and CPT-11 using HT-29-bearing nude mice

After injection of CPT-11, the concentrations of CPT-11 and SN-38 for plasma declined rapidly with time in a log-linear fashion. On the other hand, NK012 (polymer-bound SN-38) exhibited slower clearance. The clearance of NK012 in the HT-29 tumor was significantly slower and the concentration of free SN-38 was maintained at more than 30 ng/g even at 168 h after injection.

Anti-tumor activity and the distribution of NK012 and CPT-11 in SBC-3/Neo or SBC-3/VEGF tumors

In order to determine whether the potent antitumor effect of NK012 is enhanced in the tumors with high vascularity, we used vascular endothelial growth factor-secreting cells SBC-3/VEGF. There was no significant difference in the *in vitro* cytotoxic activity of each drug between SBC-3/Neo and SBC-3/VEGF. Gross findings of SBC-3/VEGF tumors are reddish as compared with SBC-3/Neo tumors. Deviating from the ordinary experimental tumor model, tumors were allowed to

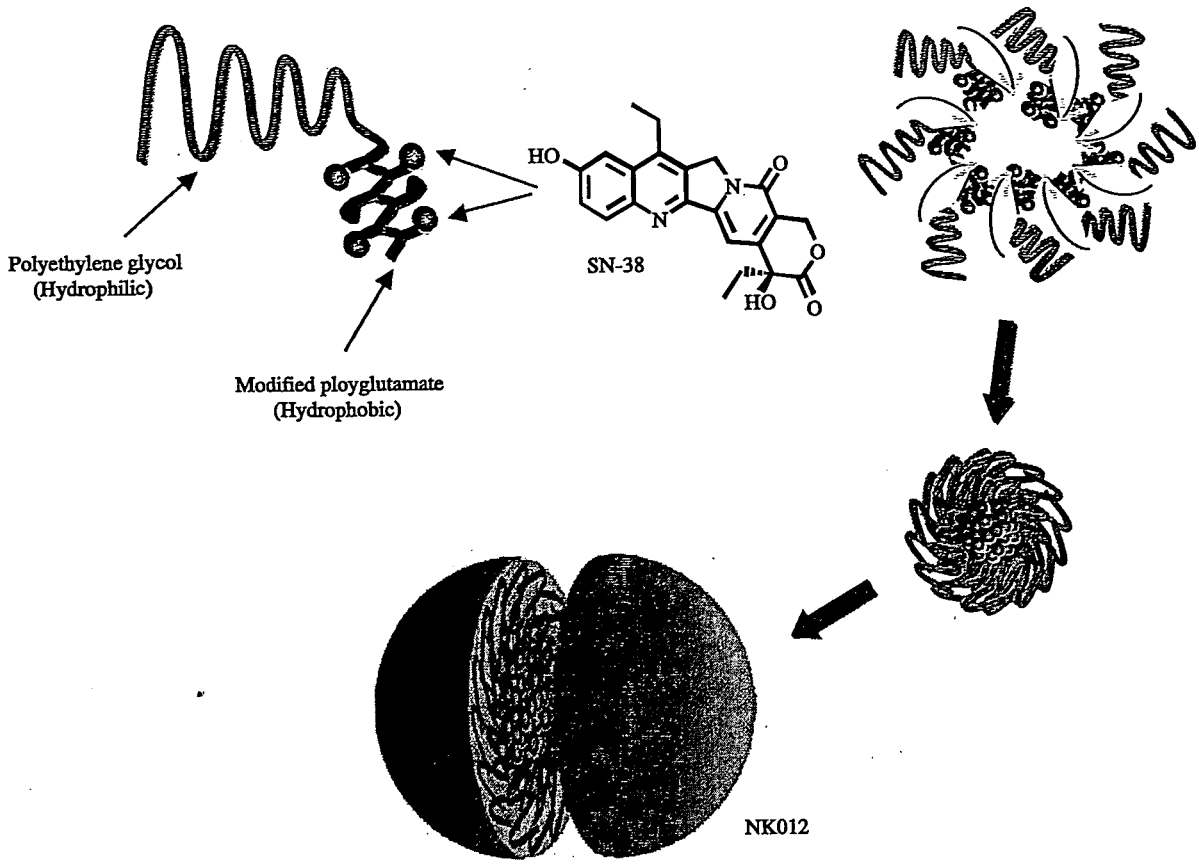

Figure 7. Schematic structure of NK012. A polymeric micelle carrier of NK012 consists of a block copolymer of PEG (molecular weight of about 5000) and partially modified polyglutamate (about 20 unit). Polyethylene glycol (hydrophilic) is believed to be the outer shell and SN-38 was incorporated into the inner core of the micelle.

grow until they became massive in size, around 1.5 cm (Figure 8), and then the treatment was initiated. NK012 at doses of 15 and 30 mg/kg showed potent anti-tumor activity against bulky SBC-3/Neo tumors (1533.1 ± 1204.7 mm^3) as compared with CPT-11 (Figure 8). Striking antitumor activity was observed in mice treated with NK012 (Figure 8) when we compared the antitumor activity of NK012 with CPT-11 using SBC-3/VEGF cells. SBC-3/VEGF bulky masses (1620.7 ± 834.0 mm^3) disappeared in all mice, although relapse 3 months after treatment was noted in one mouse treated with NK012 20 mg/kg (Figure 8). On the other hand,

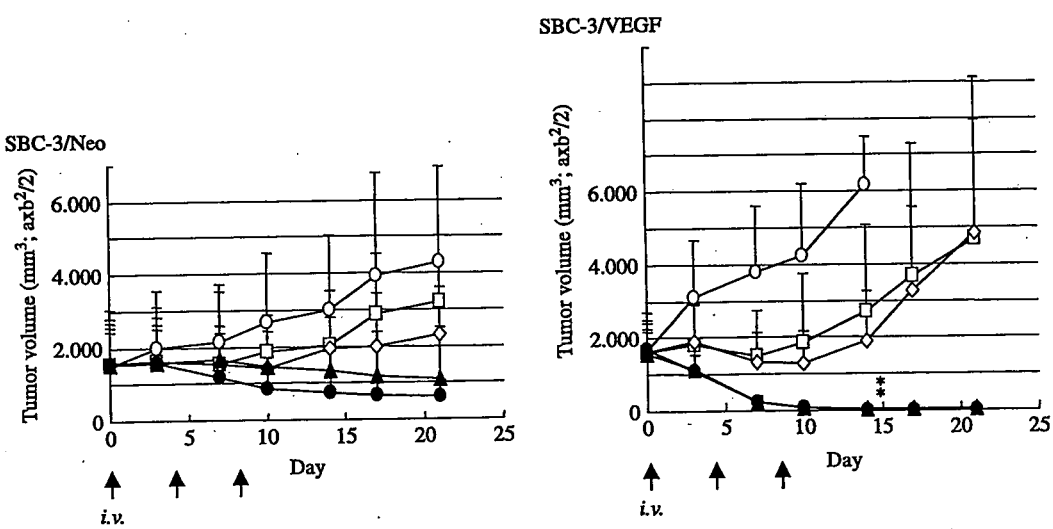

Figure 8. Intravenous administration of NK012 or CPT-11 was started when the mean tumor volumes of groups reached a massive 1500 mm^3. The mice were divided into test group (□: control; □: CPT-11 20 mg/kg/day; □: CPT-11 40 mg/kg/day; ▲: NK012 15 mg/kg/day; and □: NK012 30 mg/kg/day). NK012 or CPT-11 was administered i.v. on days 0, 4 and 8. Each group consisted of 4 mice. $*P < 0.05$.

SBC-3/VEGF were not eradicated and rapidly regrew after a partial response in mice treated with CPT-11. Approximate 10% body weight loss was observed in mice treated with NK012 20 mg/kg, but no significant difference was observed in comparison with mice treated with CPT-11 30 mg/kg.

We then examined the distribution of free SN-38 in the SBC-3/Neo and SBC-3/VEGF masses after administration of NK012 and CPT-11. In the case of CPT-11 administration, the concentrations at 1 and 6 h after the administration were less than 100 ng/g both in the SBC-3/Neo and SBC-3/VEGF tumors, and were almost negligible at 24 h in both tumors. There was no significant difference in the concentration between the SBC-3/Neo and SBC-3/VEGF tumors. On the other hand, in the case of NK012 administration, free SN-38 was detectable in the tumors even at 72 h after the administration. The concentrations of free SN-38 were higher in the SBC-3/VEGF tumors than those in the SBC-3/Neo tumors at any time point during the period of observation (significant at 1, 6, 24 h. $*P < 0.05$).

Tissue distribution of SN-38 after administration of NK012 and CPT-11

We examined the concentration–time profile of free SN-38 in various tissues after i.v. administration of NK012 and CPT-11. All organs measured exhibited the highest concentration of SN-38 at 1 h after administration in mice given CPT-11. On the other hand, mice given NK012 exhibited prolonged distribution in the liver and spleen. In a similar manner to other micellar drugs (Yokoyama et al. 1991a, b, c; Uchino et al. 2005), NK012 demonstrated relatively higher accumulation in organs of the reticuloendothelial system. In the lung, kidney and small intestine, the highest concentration of free SN-38 was achieved at 1 h after injection of NK012 and the concentration was almost negligible at 24 h. Although the concentrations of free SN-38 in the small intestine were relatively high at 1 h after administration of NK012 and CPT-11, those rapidly decreased. Interestingly, there was no significant difference in the kinetic character of free SN-38 in the small intestine between mice treated with NK012 and CPT-11.

Synergistic antitumour activity of the NK012 combined with 5-fluorouracil

In two phase III trials, the addition of CPT-11 to bolus or infusional 5FU/LV regimens clearly yielded greater efficacy than treatment with 5FU/LV alone, with a doubling of the tumor response rate and prolongation of the median survival time by 2–3 months (Douillard et al. 2000; Saltz et al. 2001).

We demonstrated that the novel SN-38-incorporating polymeric micelles, NK012, exerted superior antitumor activity and less toxicity as compared to CPT-11 (Koizumi et al. 2006). Therefore, we speculated that the use of NK012 in place of CPT-11 in combination with 5FU may also yield superior results.

Comparison of the antitumor effect of combined NK012/5FU and CPT-11/5FU. The therapeutic effect of CPT-11/5FU was apparently inferior to that of NK012/5FU or even NK012 alone at the MTD. A more potent antitumor effect, namely 100% CR rate, was obtained in the NK012 alone and NK012/5FU groups, as compared with the 0% CR rate in the CPT-11/5FU (Nakajima and Matsumura 2007).

Specificity of cell cycle perturbation. We studied the difference in the effects between NK012 and CPT-11 on the cell cycle. The data indicate that both NK012 and CPT-11 had a tendency to accumulate the 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 histograms show aneuploidy of the tumor and that administration of NK012 or CPT-11 caused apoptosis of a proportion of the tumor cells (Nakajima and Matsumura 2007).

Present situation of a clinical study of NK012

A phase 1 study of NK012 is now under way in the National Cancer Center, Tokyo and Kashiwa in patients with advanced solid tumours. NK012 is infused intravenously over 60 min every 21 days until disease progression or unacceptable toxicity occurs.

Conclusion

A quarter of a century has passed since the EPR effect was discovered (Matsumura and Maeda 1986). Now the phrase EPR has become a fundamental principle in the field of DDS. Until recently, the EPR had not been recognized in the field of oncology. However, many oncologists have now become acquainted with it, since some drugs such as doxil, abraxane, and several PEGylated proteinacious agents formulated based on the EPR have been approved in the field of oncology. Micelle carrier systems described in this chapter are obviously categorized as DDS based on the EPR. I believe that some anticancer agents incorporating micelle nanoparticles may be approved for clinical use soon.

Our next task is to develop DDS utilizing the EPR effect, which can accumulate selectively in solid tumours but also allow distribution of the delivered

bullets (anticancer agents) through the entire mass of the solid tumor tissue.

References

Bellmunt J, Ribas A, Eres N, Albanell J, Almanza C, Bermejo B, Sole LA, Baselga J. 1997. Carboplatin-based versus cisplatin-based chemotherapy in the treatment of surgically incurable advanced bladder carcinoma. Cancer 80:1966–1972.

Bodurka DC, Levenback C, Wolf JK, Gano J, Wharton JT, Kavanagh JJ, et al. 2003. Phase II trial of irinotecan in patients with metastatic epithelial ovarian cancer or peritoneal cancer. J Clin Oncol 21:291–297.

Carney DN. 1996. Chemotherapy in the management of patients with inoperable non-small cell lung cancer. Semin Oncol 23:71–75.

Cassidy J, Tabernero J, Twelves C, Brunet R, Butts C, Conroy T, Debraud F, Figer A, Grossmann J, Sawada N, Schoffski P, Sobrero A, Van Cutsem E, Diaz-Rubio E. 2004. XELOX (capecitabine plus oxaliplatin): Active first-line therapy for patients with metastatic colorectal cancer. J Clin Oncol 22: 2084–2091.

Cleare MJ, Hydes PC, Malerbi BW, Watkins DM. 1978. Anti-tumor platinum complexes: Relationships between chemical properties and activity. Biochimie 60:835–850.

Cunningham D, Pyrhonen S, James RD, Punt CJ, Hickish TF, Heikkila R, et al. 1998. Randomised trial of irinotecan plus supportive care versus supportive care alone after fluorouracil failure for patients with metastatic colorectal cancer. Lancet 352: 1413–1418.

Douillard JY, Cunningham D, Roth AD, et al. 2000. Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: A multicentre randomised trial. Lancet 355:1041–1047.

du Bois A, Luck HJ, Meier W, Adams HP, Mobus V, Costa S, Bauknecht T, Richter B, Warm M, Schroder W, Olbricht S, Nitz U, Jackisch C, Emons G, Wagner U, Kuhn W, Pfisterer J. 2003. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. J Natl Cancer Inst 95:1320–1329.

Gallo RC, Whang-Peng J, Adamson RH. 1971. Studies on the antitumor activity, mechanism of action, and cell cycle effects of camptothecin. J Natl Cancer Inst 46:789–795.

Gottlieb JA, Guarino AM, Call JB, Oliverio VT, Block JB. 1970. Preliminary pharmacologic and clinical evaluation of camptothecin sodium (NSC-100880). Cancer Chemother Rep 54: 461–470.

Guichard S, Terret C, Hennebelle I, Lochon I, Chevreau P, Fretigny E, et al. 1999. CPT-11 converting carboxylesterase and topoisomerase activities in tumour and normal colon and liver tissues. Br J Cancer 80:364–370.

Horwich A, Sleijfer DT, Fossa SD, Kaye SB, Oliver RT, Cullen MH, Mead GM, de Wit R, de Mulder PH, Dearnaley DP, Cook PA, Sylvester RJ, Stenning SP. 1997. Randomized trial of bleomycin, etoposide, and cisplatin compared with bleomycin, etoposide, and carboplatin in good-prognosis metastatic nonseminomatous germ cell cancer: A Multiinstitutional Medical Research Council/European Organization for Research and Treatment of Cancer Trial. J Clin Oncol 15:1844–1852.

Horwich A, Sleijfer DT, Fossa SD, Kaye SB, Oliver RT, Cullen MH, Mead GM, de Wit R, de Mulder PH, Dearnaley DP, Cook PA, Sylvester RJ, Stenning SP. 1997. Randomized trial of bleomycin, etoposide, and cisplatin compared with bleomycin, etoposide, and carboplatin in good-prognosis metastatic nonseminomatous germ cell cancer: A Multiinstitutional Medical Research Council/European Organization for Research and Treatment of Cancer Trial. J Clin Oncol 15:1844–1852.

Kataoka K, Kwon GS, Yokoyama M, Okano T, Sakurai Y. 1993. Block copolymer micelles as vehicles for drug delivey. J Controlled Release 24:119–132.

Kato K, Hamaguchi T, Matsumura Y, Yasui H, Okusaka T, Ueno H, Ikeda M, Muro K, Shirao K, Shimada Y. 2018. Phase I study of NK105, polymer micelle paclitaxel, in patients with advanced cancer. Proc Am Soc Clin Oncol 25:2006.

Khayat D, Antoine EC, Coeffic D. 2000. Taxol in the management of cancers of the breast and the ovary. Cancer Invest 18: 242–260.

Koizumi F, Kitagawa M, Negishi T, et al. 2006. Novel SN-38-incorporating polymeric micelles, NK012, eradicate vascular endothelial growth factor-secreting bulky tumors. Cancer Res 66:10048–10056.

Li LH, Fraser TJ, Olin EJ, Bhuyan BK. 1972. Action of camptothecin on mammalian cells in culture. Cancer Res 32: 2643–2650.

Matsumura Y, Maeda H. 1986. A new concept for macromolecular therapeutics in cancer chemotherapy: Mechanism of tumoritropic accumulation of proteins and the antitumor agent smancs. Cancer Res 46:6387–6392.

Muggia FM, Creaven PJ, Hansen HH, Cohen MH, Selawry OS. 1972. Phase I clinical trial of weekly and daily treatment with camptothecin (NSC-100880): Correlation with preclinical studies. Cancer Chemother Rep 56:515–521.

Nakajima T, Matsumura Y. 2007. Abstr 4729 Synergictic antitumor activity of novel polymeric micelles incorporating SN-38(NK102) in combination with 5-fluorouracil (5-FU) in mouse model of colon cancer. Proc Am Assoc Cancer Res 98.

Negoro S, Masuda N, Takada Y, Sugiura T, Kudoh S, Katakami N, et al. 2003. CPT-11 Lung Cancer Study Group West. Randomised phase III trial of irinotecan combined with cisplatin for advanced non-small-cell lung cancer. Br J Cancer 88: 335–341.

Nishiyama N, Okazaki S, Cabral H, Miyamoto M, Kato Y, Sugiyama Y, Nishio K, Matsumura Y, Kataoka K. 2003. Novel cisplatin-incorporated polymeric micelles can eradicate solid tumors in mice. Cancer Res 63:8977–8983.

Noda K, Nishiwaki Y, Kawahara M, Negoro S, Sugiura T, Yokoyama A, et al. 2002. Irinotecan plus cisplatin compared with etoposide plus cisplatin for extensive small-cell lung cancer. N Engl J Med 346:85–91.

Pinzani V, Bressolle F, Haug IJ, Galtier M, Blayac JP, Balmes P. 1994. Cisplatin-induced renal toxicity and toxicity-modulating strategies: A review. Cancer Chemother Pharmacol 35:1–9.

Roth BJ. 1996. Chemotherapy for advanced bladder cancer. Semin Oncol 23:633–644, Screnci D, McKeage MJ, Galettis P, Hambley TW, Palmer BD and Baguley BC (2000). Relationships between hydrophobicity, reactivity, accumulation and peripheral nerve toxicity of a series of platinum drugs. Br J Cancer, 82, 966–72.

Rothenberg ML, Kuhn JG, Burris HA, 3rd, Nelson J, Eckardt JR, Tristan-Morales M, et al. 1993. Phase I and pharmacokinetic trial of weekly CPT-11. J Clin Oncol 11:2194–2204.

Rowinsky EK, Donehower RC. 1995. Paclitaxel (taxol). N Engl J Med 332:1004–1014, Savic R, Luo L, Eisenberg A, Maysinger D (2003). Micellar nanocontainers distribute to defined cytoplasmic organelles. Science, 300, 615–8.

Rowinsky EK, Chaudhry V, Forastiere AA, Sartorius SE, Ettinger DS, Grochow LB, Lubejko BG, Cornblath DR, Donehower RC. 1993. Phase I and pharmacologic study of paclitaxel and cisplatin with granulocyte colony-stimulating factor: Neuromuscular toxicity is dose-limiting. J Clin Oncol 11:2010–2020.

Saltz LB, Cox JV, Blanke C, Rosen LS, Fehrenbacher L, Moore MJ, et al. 2000. Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. N Engl J Med 343:905–914.

Saltz LB, Douillard JY, Pirotta N, et al. 2001. Irinotecan plus fluorouracil/leucovorin for metastatic colorectal cancer: A new survival standard. Oncologist 6:81–91.

Slatter JG, Schaaf LJ, Sams JP, Feenstra KL, Johnson MG, Bombardt PA, et al. 2000. Pharmacokinetics, metabolism, and excretion of irinotecan (CPT-11) following I.V. infusion of [(14)C]CPT-11 in cancer patients. Drug Metab Dispos 28: 423–433.

Takimoto CH, Arbuck SG. 2001. Topoisomerase I targeting agents: The camptothecins. In: Chabner BA, Lango DL, editors. Cancer chemotherapy and biotherapy: Principal and practice. 3rd ed., Philadelphia (PA): Lippincott Williams & Wilkins. p 579–646.

Uchino H, Matsumura Y, Negishi T, Koizumi F, Hayashi T, Honda T, et al. 2005. Cisplatin-incorporating polymeric micelles (NC-6004) can reduce nephrotoxicity and neurotoxicity of cisplatin in rats. Br J Cancer 93:678–687.

Wasserheit C, Frazein A, Oratz R, Sorich J, Downey A, Hochster H, Chachoua A, Wernz J, Zeleniuch-Jacquotte A, Blum R, Speyer J. 1996. Phase II trial of paclitaxel and cisplatin in women with advanced breast cancer: An active regimen with limiting neurotoxicity. J Clin Oncol 14:1993–1999.

Weiss RB, Donehower RC, Wiernik PH, Ohnuma T, Gralla RJ, Trump DL, Baker JR, Jr, Van Echo DA, Von Hoff DD, Leyland-Jones B. 1990. Hypersensitivity reactions from taxol. J Clin Oncol 8:1263–1268.

Yokoyama M, Miyauchi M, Yamada N, Okano T, Sakurai Y, Kataoka K, Inoue S. 1990. Polymer micelles as novel drug carrier: Adriamycin-conjugted poly(ethylene glycol)-poly(aspartic acid) block copolymer. J Controlled Release 11:269–278.

Yokoyama M, Okano T, Sakurai Y, Ekimoto H, Shibazaki C, Kataoka K. 1991. Toxicity and antitumour activity against solid tumours of micelle-forming polymeric anticancer drug and its extremely long circulation in blood. Cancer Res 51:3229–3236.

Yokoyama M, Okano T, Sakurai Y, Ekimoto H, Shibazaki C, Kataoka K. 1991. Toxicity and antitumour activity against solid tumours of micelle-forming polymeric anticancer drug and its extremely long circulation in blood. Cancer Res 51:3229–3236.

Yokoyama M, Okano T, Sakurai Y, Ekimoto H, Shibazaki C, Kataoka K. 1991. Toxicity and antitumor activity against solid tumors of micelle-forming polymeric anticancer drug and its extremely long circulation in blood. Cancer Res 51:3229–3236.

J Hum Genet (2007) 52:804–819
DOI 10.1007/s10038-007-0186-6

ORIGINAL ARTICLE

Genetic variations and haplotype structures of the *DPYD* gene encoding dihydropyrimidine dehydrogenase in Japanese and their ethnic differences

Keiko Maekawa · Mayumi Saeki · Yoshiro Saito · Shogo Ozawa · Kouichi Kurose · Nahoko Kaniwa · Manabu Kawamoto · Naoyuki Kamatani · Ken Kato · Tetsuya Hamaguchi · Yasuhide Yamada · Kuniaki Shirao · Yasuhiro Shimada · Manabu Muto · Toshihiko Doi · Atsushi Ohtsu · Teruhiko Yoshida · Yasuhiro Matsumura · Nagahiro Saijo · Jun-ichi Sawada

Received: 30 May 2007 / Accepted: 26 July 2007 / Published online: 9 September 2007

Abstract Dihydropyrimidine dehydrogenase (DPD) is an inactivating and rate-limiting enzyme for 5-fluorouracil (5-FU), and its deficiency is associated with a risk for developing a severe or fatal toxicity to 5-FU. In this study, to search for genetic variations of *DPYD* encoding DPD in Japanese, the putative promoter region, all exons, and flanking introns of *DPYD* were sequenced from 341 subjects including cancer patients treated with 5-FU. Fifty-five genetic variations, including 38 novel ones, were found and consisted of 4 in the 5′-flanking region, 21 (5 synonymous and 16 nonsynonymous) in the coding exons, and 30 in the introns. Nine novel nonsynonymous SNPs, 29C>A (Ala10Glu), 325T>A (Tyr109Asn), 451A>G (Asn151Asp), 733A>T (Ile245Phe), 793G>A (Glu265Lys), 1543G>A (Val515Ile), 1572T>G (Phe524Leu), 1666A>C (Ser556-Arg), and 2678A>G (Asn893Ser), were found at allele frequencies between 0.15 and 0.88%. Two known nonsynonymous variations reported only in Japanese, 1003G>T (**11*, Val335Leu) and 2303C>A (Thr768Lys), were found at allele frequencies of 0.15 and 2.8%, respectively. SNP and haplotype distributions in Japanese were quite different from those reported previously in Caucasians. This study provides fundamental information for pharmacogenetic studies for evaluating the efficacy and toxicity of 5-FU in Japanese and probably East Asians.

Keywords *DPYD* · SNP · Haplotype · Japanese · 5-fluorouracil

K. Maekawa (✉) · Y. Saito · J. Sawada
Division of Biochemistry and Immunochemistry,
National Institute of Health Sciences,
1-18-1 Kamiyoga, Setagaya-ku,
Tokyo 158-8501, Japan
e-mail: maekawa@nihs.go.jp

K. Maekawa · M. Saeki · Y. Saito · S. Ozawa · K. Kurose · N. Kaniwa · J. Sawada
Project Team for Pharmacogenetics,
National Institute of Health Sciences, Tokyo, Japan

S. Ozawa
Division of Pharmacology,
National Institute of Health Sciences, Tokyo, Japan

K. Kurose · N. Kaniwa
Division of Medicinal Safety Science,
National Institute of Health Sciences, Tokyo, Japan

M. Kawamoto · N. Kamatani
Division of Genomic Medicine,
Department of Advanced Biomedical Engineering and Science,
Tokyo Women's Medical University, Tokyo, Japan

K. Kato · T. Hamaguchi · Y. Yamada · K. Shirao · Y. Shimada
Gastrointestinal Oncology Division, National Cancer Center Hospital, National Cancer Center, Tokyo, Japan

M. Muto
Gastrointestinal Oncology Division,
National Cancer Center Hospital East, Kashiwa, Japan

T. Doi · A. Ohtsu
Division of GI Oncology/Digestive Endoscopy,
National Cancer Center Hospital East, Kashiwa, Japan

T. Yoshida
Genetics Division, National Cancer Center Research Institute,
National Cancer Center, Tokyo, Japan

Y. Matsumura
Research Center of Innovative Oncology,
National Cancer Center Hospital East, Kashiwa, Japan

N. Saijo
Deputy Director, National Cancer Center Hospital East, Kashiwa, Japan

Introduction

Dihydropyrimidine dehydrogenase (DPD) is an inactivating and rate-limiting enzyme for 5-fluorouracil (5-FU), which is used in various therapeutic regimens for gastrointestinal, breast and head/neck cancers (Grem 1996). While the antitumor effect of 5-FU is exerted via anabolic pathways responsible for its intracellular conversion into anti-proliferative nucleotides, DPD affects 5-FU availability by rapidly degrading it to 5, 6-dihydrofluorouracil (DHFU) (Heggie et al. 1987). The importance of DPD in 5-FU metabolism was also highlighted by a lethal drug interaction between 5-FU and the antiviral agent sorivudine. Due to inhibition of DPD by a sorivudine metabolite, severe systemic exposure to 5-FU caused several acute deaths in Japan (Nishiyama et al. 2000).

5-FU catabolism occurs in various tissues, including tumors, but is highest in the liver (Naguib et al. 1985; Lu et al. 1993). Wide variations in DPD activity (8- to 21-fold) were shown in Caucasians, and 3–5% of Caucasians had reduced DPD activity (Etienne et al. 1994; Lu et al. 1998). This variability, which is partially attributed to genetic defects of the DPD gene (*DPYD*), leads to differential responses of cancer patients, resistance to or increased toxicity of 5-FU (van Kuilenburg 2004). Complete DPD deficiency is also associated with the inherited metabolic disorder, thymine-uraciluria, which is characterized by neurological problems in pediatric patients (Bakkeren et al. 1984).

To date, at least 30 variant *DPYD* alleles have been published, with or without deleterious impact upon DPD activity (Gross et al. 2003; Ogura et al. 2005; Seck et al. 2005; van Kuilenburg 2004; Zhu et al. 2004). Of these variations, a splice site polymorphism, IVS14 + 1G>A, which causes skipping of exon 14, is occasionally detected in North Europeans with allele frequencies of 0.01–0.02 (van Kuilenburg 2004). Detection of IVS14 + 1G>A in patients suffering from 5-FU-associated grade 3 or 4 toxicity revealed that 24–28% of them were heterozygous or homozygous for this single nucleotide polymorphism (SNP) (van Kuilenburg 2004). However, this SNP has not been reported in Japanese and African-Americans. Recently, Ogura et al. (2005) have shown that a Japanese population exhibits a large degree of interindividual variations in DPD activity of peripheral blood mononuclear cells. They also identified a novel variation, 1097G>C (Gly366Ala), in a healthy volunteer with the lowest DPD activity and demonstrated that the 366Ala variant has reduced activity towards 5-FU in vitro. At present, however, information on variant alleles with clinical relevance in Japanese is limited and cannot fully explain polymorphic DPD activity.

In this study, we searched for genetic variations in *DPYD* by sequencing 5′ regulatory regions, all exons and surrounding introns from 341 Japanese subjects. Fifty-five variations including nine novel nonsynonymous ones were identified. Then, linkage disequilibrium (LD) and haplotype analyses were performed to clarify the *DPYD* haplotype structures in Japanese.

Materials and methods

Human DNA samples

Three hundred and forty-one Japanese subjects in this study included 263 cancer patients and 78 healthy volunteers. All 263 patients were administered 5-FU or tegafur for treatment of various cancers (mainly stomach and colon) at the National Cancer Center, and blood samples were collected prior to the fluoropyrimidine chemotherapy. The healthy volunteers were recruited at the Tokyo Women's Medical University. DNA was extracted from the blood of cancer patients and Epstein-Barr virus-transformed lymphoblastoid cells derived from healthy volunteers. Written informed consent was obtained from all participating subjects. The ethical review boards of the National Cancer Center, the Tokyo Women's Medical University and the National Institute of Health Sciences approved this study.

PCR conditions for DNA sequencing

To amplify 22 exons (exons 2–23) of *DPYD*, multiplex PCRs were performed by using four sets of mixed primers (mix 1 to mix 4 of "first PCR" in Table 1). Namely, five exonic fragments were simultaneously amplified from 50 ng of genomic DNA using 0.625 units of Ex-Taq (Takara Bio. Inc., Shiga, Japan) with 0.20 μM primers. Because of the high GC content in exon 1 of *DPYD*, this region was separately amplified from 50 ng of genomic DNA with 2.5 units of LA-Taq and 0.2 μM primers (listed in Table 1) in GC buffer I (Takara Bio. Inc.). The first PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 58°C for 1 min, and 72°C for 2 min; and then a final extension for 7 min at 72°C. Next, each exon was amplified separately from the first PCR products by nested PCR (2nd PCR) using the primer sets (0.2 μM) listed in "second PCR" of Table 1. The second PCR conditions were the same as those of the first PCR, and LA-Taq (2.5 units) for exon 1 and Ex-Taq (0.625 units) for exons 2–23 were used. All PCR primers were designed in the flanking intronic sites to analyze the exon-intron splice junctions. The PCR products were treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH) and sequenced directly on both strands using an ABI BigDye Terminator Cycle Sequencing Kit (Applied Biosystems,

Table 1 Primer sequences for human *DPYD*

Amplified and sequenced region			Forward primer		Reverse primer		PCR product (bp)
			Sequences (5' to 3')	Position[a]	Sequences (5' to 3')	Position[a]	
First PCR	5'-UTR to exon 1		GTTCTGGAAGGTAATCTGATGG	52207178	ACGACATACAGGAGGTGAAG	52205443	1,736
	Mix 1	Exon 2	CTACTTGGGAGACTAAGGTG	52168526	GTATCATTGTGTCATTAGGC	52167832	695
		Exon 3	TCCCTTCATCTTAGTCAATG	52113605	CTGAGGCTTAACATTTATGC	52112876	730
		Exon 4	TCTGAGAGGAGGGACAGTTA	52025660	AATCACAACTTGGAAGTGCT	52025165	496
		Exon 5	AAATGGAGGATAACCTGAGT	52007046	TAATAACCTGCTGGGATTGC	52006234	813
		Exon 6	AGAGGAGAGGCACTTAATGT	51984772	TGCTTCAAGCCAACTGCAAA	51984115	658
	Mix 2	Exon 7	CTCAAATAATAGTGCCATAGG	51977410	CAGTAGACAGACAAATGCCC	51976498	913
		Exon 8	CACATCGTGCTTTGAACATA	51964415	CCAACTCCATCCTTTATGAT	51963667	749
		Exons 9 and 10	TGAGGCAAGAATATAACCTG	51880431	TCCGTATGTGTCTTATTACC	51877795	2,637
		Exon 11	AGAAATACCTTATGATGCCG	51859160	GCCTTTTGAATCAAGATTGC	51858562	599
		Exon 16	CTCCCTATGCTTCAGTTCAC	51658925	TGCCGTGCCCCATTTACTAC	51658114	812
	Mix 3	Exon 12	CCGCTCTGAAACATTGACCA	51834944	CTGGGATTATAGGCATTAGG	51834279	666
		Exon 13	GCCCATATCTCTGAGCACTA	51801258	ATCTTTGTTGCTTCCTAGAC	51800450	809
		Exon 14	CCTTCACTGATTTACATCGG	51735640	CCAGCCACATACAGTGAAAA	51734704	937
		Exon 15	AGCCAGTAAAATCCTCTCTA	51667711	TATGGAAAACCTGCTGACTA	51666815	897
		Exon 23	TGGAAAGACCCGAACTCTGC	51364409	AGCGAAGGGGATTTTACTTA	51363336	1,074
	Mix 4	Exons 17 and 18	TTCTAAAGGCTCTGTTGAGG	51591491	TGGCAAAAGAACTGAGAGAC	51589933	1,559
		Exon 19	CGTGGATTCAAGCAGTTTTC	51520500	AGACAGTGGGTTCGTAAGCC	51519586	915
		Exon 20	CTGTGACACCATTACCATTG	51478435	TGCCAGTCATCACCACAGTA	51477733	703
		Exon 21	GAACCTGATACCGAGAAGAC	51383758	AAATGTCCAGGCTTTCCAGA	51382987	772
		Exon 22	GCCATAACAACTCACACGGG	51367740	TTGGCAGAAGGAATCATAGC	51366885	856
Second PCR		5'-UTR to exon 1	TGTGGATGTTTTTGCTCGC	52206503	AGTAAACAGGTCCCGACGC	52205586	918
		Exon 2	GTGAACTGAGATTGTACCACTGC	52168471	CATATCCCTTATCAAAAATGCTT	52167924	548
		Exon 3	GAATGCTACCCAATTAAAGTGG	52113285	TTCAAAACCAAATACAGCCTC	52112899	387
		Exon 4	TGCCAAAGATGAAACACAGA	52025601	ACCCACAGATAATAGAGAACAAGA	52025273	329
		Exon 5	TGATGGTTCCTGATAGTAGTATTG	52006775	TGTCACACTAAAAATGTTGGG	52006348	428
		Exon 6	AAGGAAAGACTGAAAGTTAGCC	51984688	GAGCCTGAAGTTCCTATATGAT	51984201	488
		Exon 7	TTCTACTGTATCTTCACTCCACG	51976953	GCTTCTGCCTGATGTAGC	51976541	413
		Exon 8	GGCTGACTTTTCATTCTTTTT	51964221	CATCTTGCCGAAATCTCTCC	51963831	391
		Exon 9	TGTGATTTACGATGTGTACTTGG	51880335	GCAAGGTTGGGTGTGAGAG	51879895	441
		Exon 10	AAAATGGGAATAAAACTGTCTT	51878507	TCAGGATATGGAAGACTTAGCAC	51877859	649
		Exon 11	ACTGGTAACTGAAACTCAG	51859069	CAATTCCCTGAAAGCTAG	51858628	442
		Exon 12	TCAGTGCCCTTCAAATGTGT	51834881	ACCAAATAGAAATGCTCTTATAGA	51834414	468
		Exon 13	TCGGATGCTGTGTTGAAGTG	51800982	TGTGTAATGATAGGTCGTGTC	51800543	440

Table 1 Primer sequences for human *DPYD*

Amplified and sequenced region		Forward primer		Reverse primer		PCR product (bp)
		Sequences (5' to 3')	Position[a]	Sequences (5' to 3')	Position[a]	
	Exon 14	TGCAAATATGTGAGGAGGGACC	51735287	CAGCAAAGCAACTGGCAGATTC	51734877	411
	Exon 15	GCTATCTTACCCTGCTATTTTC	51667571	TAGGTAGTGTGTGAAATCCAAGG	51667107	465
	Exon 16	CCCCTTATGAGCACTGAGTAAAT	51658821	TAGTAACTATCCATACGGGGG	51658440	382
	Exon 17	AGTCTAGGTGTAATACTGAGGAGG	51591407	ATCAAGTGCTCAACTGGAAACT	51590986	422
	Exon 18	GTGAAGAACTTTGAGGAGAAGAC	51590461	CATCCTGTGCTGTCACTTGA	51590026	436
	Exon 19	ATTTGTCCAGTGACGCTGTC	51520048	TCAGGTCTCTTCATAACTTGTCAG	51519629	420
	Exon 20	GAGAAGTGAATTTGTTTGGAG	51478265	TTTGTTAGTGAGAATGTGAGATGG	51477926	340
	Exon 21	AGTGGTCCAAAACAATGAGTG	51383737	TGCTTGCCAGTGTTCTAAAA	51383221	517
	Exon 22	GGGTGTCATTTATTCTTTCTGTC	51367723	GGCTGATGAAATGGTATAAAAA	51367033	691
	Exon 23	GTTGTCTCATAGTGTGGCTCCTC	51364206	TTTTTCACATAAGACAACTGGCA	51363641	566
Sequencing	5'-UTR to exon 1	TGTGGATGTTTTTGCTCGC	52206503			
	5'-UTR to exon 1	CGGACTGCTTTTACCTTTGC	52206258	CCAGAGAGCCAAGTGACAGC	52205933	
	5'-UTR to exon 1	CCCTAGTCTGCCTGTTTTCG	52205987	AGTAAACAGGTCCCGACGC	52205586	
	Exon 2	GTGACAAAGTGAGAGAGACCGT	52168436	GCCTTACAATGTGTGGAGTGAG	52168152	
	Exon 3	GAATGCTACCCAATTAAAGTGG	52113285	TTCAAAACCAAATACAGCCTC	52112899	
	Exon 4	TGCCAAAGATGAAACACAGA	52025601	ACCCACAGATAATAGAGAACAAGA	52025273	
	Exon 5	TGATGGTTCCTGATAGTAGTATTG	52006775	TGTCACACTAAAAATGTTGGG	52006348	
	Exon 6	AAAATATGTTTGAGGATGTAAGC	51984560	GAGCCTGAAGTTCCTATATGAT	51984201	
	Exon 7	TTCTACTGTATCTTCACTCCACG	51976953	GCTTCTGCCTGATGTAGC	51976541	
	Exon 8	GGCTGACTTTTCATTCTTTTT	51964221	CATCTTGCCGAAATCTCTCC	51963831	
	Exon 9	TGTGATTTACGATGTGTACTTGG	51880335	GCAAGGTTGGGTGTGAGAG	51879895	
	Exon 10	AAAATGGGAATAAAACTGTCTT	51878507	TTCATCTCCTAAAATCTGTTGG	51878109	
	Exon 11	ACTGGTAACTGAAACTCAG	51859069	CAATTCCCTGAAAGCTAG	51858628	
	Exon 12	TCAGTGCCCTTCAAATGTGT	51834881	GAGTATCAAAAATAAATGAAGCAC	51834439	
	Exon 13	TCGGATGCTGTGTTGAAGTG	51800982	TGTGTAATGATAGGTCGTGTC	51800543	
	Exon 14	TGCAAATATGTGAGGAGGGACC	51735287	CAGCAAAGCAACTGGCAGATTC	51734877	
	Exon 15	GCTATCTTACCCTGCTATTTTC	51667571	TAGGTAGTGTGTGAAATCCAAGG	51667107	
	Exon 16	CCCCTTATGAGCACTGAGTAAAT	51658821	TAGTAACTATCCATACGGGGG	51658440	
	Exon 17	AGTCTAGGTGTAATACTGAGGAGG	51591407	ATCAAGTGCTCAACTGGAAACT	51590986	
	Exon 18	GTGAAGAACTTTGAGGAGAAGAC	51580461	CATCCTGTGCTGTCACTTGA	51590026	
	Exon 19	ATTTGTCCAGTGACGCTGTC	51520048	CGAATCTATTTTTTTTTTGTCAC	51519715	
	Exon 20	GAGAAGTGAATTTGTTTGGAG	51478265	TTTGTTAGTGAGAATGTGAGATGG	51477926	
	Exon 21	TATCTTCCCATTTTTCTCTTCTC	51383644	TGCCAGTGTTCTAAAAAGTATAAA	51383225	
	Exon 22	GTATAAAAACAGGAAAATGCTGA	51367510	ATAAGGGTGACAGGACAGAAAG	51367125	
	Exon 23	GTTGTCTCATAGTGTGGCTCCTC	51364206	TATTTGTTTTAATTTGGAAAGAG	51363821	

[a] Nucleotide position of the 5' end of each primer on NT_032977.7

Foster City, CA) with the primers listed in "sequencing" of Table 1. Excess dye was removed with a DyeEx96 kit (Qiagen, Hilden, Germany). The eluates were analyzed on an ABI Prism 3700 DNA Analyzer (Applied Biosystems). All novel SNPs were confirmed by sequencing of PCR products generated from new genomic DNA amplifications. The genomic and cDNA sequences of *DPYD* obtained from GenBank (NT_032977.7 and NM_000110.2, respectively) were used as reference sequences. SNP positions were numbered based on the cDNA sequence, and adenine of the translational initiation site in exon 1 was numbered +1. For intronic polymorphisms, the position was numbered from the nearest exon.

Linkage disequilibrium (LD) and haplotype analyses

Hardy-Weinberg equilibrium and LD analyses were performed by SNPAlyze software (Dynacom Co., Yokohama, Japan), and pairwise LD parameters between variations were obtained as the |D'| and rho square (r^2) values. Some haplotypes were unambiguously identified from subjects with homozygous variations at all sites or a heterozygous variation at only one site. Diplotype configurations were inferred by LDSUPPORT software, which determines the posterior probability distribution of the diplotype for each subject based on the estimated haplotype frequencies (Kitamura et al. 2002). Although the nomenclature for nonsynonymous *DPYD* alleles (*DPYD*1* to *DPYD*13*) have been already publicized (McLeod et al. 1998; Collie-Duguid et al. 2000; Johnson et al. 2002), several reported alleles remain unassigned. To avoid confusion with the previous *DPYD* allele nomenclature, our block haplotypes in this study were tentatively defined by using '#' instead of '*'. A group of haplotypes without any amino acid change is designated as *#1*, and the haplotype groups bearing already defined alleles, *DPYD*5* (Ile543Val), *DPYD*6* (Val732Ile), *DPYD*9* (Cys29Arg) and *DPYD*11* (Val335Leu), were numbered by using the corresponding Arabic numerals, *#5*, *#6*, *#9*, and *#11*, respectively. Other haplotypes with known nonsynonymous SNPs such as 496A>G (Met166Val) or with the novel nonsynonymous SNP were represented by '#' plus amino acid positions followed by variant residues (for example, *#166V*). Subtypes within each haplotype group were consecutively named with small alphabetical letters depending on their frequencies. Haplotypes ambiguously inferred in only one patient were indicated in the Fig. 3 legend. Combinations of block haplotypes were analyzed by Haploview software (http://www.broad.mit.edu/mpg/haploview/index.php) (Barrett et al. 2005), and the long-range (whole gene) haplotypes spanning all blocks were inferred by Hapblock software (www.cmb.usc.edu/msms/HapBlock/) (Zhang et al. 2005).

Typing data on *DPYD* from unrelated 44 Japanese and 30 Caucasian trios were also obtained from the HapMap project (HapMap release 19: http://www.hapmap.org/). The LD profiles and haplotypes of the HapMap data were obtained by Marker beta in Gmap Net (http://www.gmap.net/marker) using its four (1254711, 1254712, 1254713, and 1254714) and six (1166276, 1166277, 1166278, 1166279, 1166280, and 1166281) datasets covering *DPYD* genomic regions for Japanese and Caucasians, respectively.

Drawing of protein structures

The coordinate data (1gth) of the crystal structure of pig DPD (Dobritzsch et al. 2002) was obtained from the Protein Data Bank. Protein Explorer (http://proteinexplorer.org) (Martz 2002) was used to display the structural features of pig DPD and depict three-dimensional views.

Results

DPYD variations found in a Japanese population

We identified 55 variations, including 38 novel ones by sequencing the promoter regions (up to 613 bp upstream from the translational initiation site), all 23 exons and their flanking regions of *DPYD* from 341 Japanese subjects (Table 2). The distribution of the variations consisted of 4 in the 5' flanking region, 21 (5 synonymous and 16 nonsynonymous ones) in the coding exons (Fig. 1) and 30 in the introns. Since we did not find any significant differences in allele frequencies between healthy volunteers and cancer patients ($P > 0.05$ by χ^2 test or Fisher's exact test) except for one variation, IVS14 + 19C>A, ($P = 0.027$ by Fisher's exact test); the data for all subjects were analyzed as one group. All detected variations except for 451A>G (Asn151Asp) and IVS13 + 40G>A were in Hardy-Weinberg equilibrium ($P \geq 0.24$).

Thirteen novel variations in the coding region (enclosed by a square in Fig. 1) contain four synonymous SNPs, 474T>C (Phe158Phe), 639C>T (Asp213Asp), 1752A>G (Thr584Thr), and 2424T>C (Ser808Ser) and nine nonsynonymous SNPs, 29C>A (Ala10Glu), 325T>A (Tyr109Asn), 451A>G (Asn151Asp), 733A>T (Ile245Phe), 793G>A (Glu265Lys), 1543G>A (Val515Ile), 1572T>G (Phe524Leu), 1666A>C (Ser556Arg), and 2678A>G (Asn893Ser). 451A>G (Asn151Asp), 325T>A (Tyr109Asn), and 2678A>G (Asn893Ser) were found at frequencies of 0.009, 0.003 and 0.003, respectively. The others were detected as single heterozygotes (allele frequencies = 0.0015).

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/TTCCCCTCCCCGC			0.021
MPJ6_DPD002[a]		5′-Flank	52206348	–477	TTGAGGAGTTCCT/GGAAAATGCAGTT			0.026
MPJ6_DPD003[a]		5′-Flank	52206137	–266	CTCCCTCCCTCCC/ATTCTGCTTGCAG			0.045
MPJ6_DPD004[a]		5′-Flank	52206114	–243	AGGCTGGGGCGCG/AGAGCGGGCTGAA			0.0059
MPJ6_DPD005[a]		Exon 1	52205843	29	GTAAGGACTCGGC/AGGACATCGAGGT	Ala10Glu		0.0015
MPJ6_DPD006[g]	rs1801265	Exon 2	52168278	85	CATGCAACTCTGT/CGTTCCACTTCGG	Cys29Arg	*9	0.029
MPJ6_DPD007[a]		Intron 2	52168055	IVS2 + 158	TTTGAAAGTGTAT/CTTTTAATTACAC			0.0015
MPJ6_DPD008[a]		Intron 3	52113040	IVS3 + 23	GTCACCATAGCAA/GCAGTCACAGATG			0.0029
MPJ6_DPD009[a]		Exon 5	52006617	325	ATTTTGCAGAACT/AATTATGGAGCTG	Tyr109Asn		0.0029
MPJ6_DPD010[a]		Exon 5	52006491	451	GAGGGACCCATTA/GATATTGGTGGAT	Asn151Asp		0.0088
MPJ6_DPD011[a]		Exon 5	52006468	474	ATTGCAGCAATTT/CGCTACTGAGGTA	Phe158Phe		0.0044
MPJ6_DPD012[a]		Intron 5	51984611	IVS5–115	CATATTAATACTG/AAAAATGTACTGC			0.021
MPJ6_DPD013[g]	rs2297595	Exon 6	51984484	496	GTATTCAAAGCAA/GTGAGTATCCCAC	Met166Val		0.022
MPJ6_DPD014[a]		Exon 6	51984341	639	GGGGTACTCTGAC/TATCACTATATTT	Asp213Asp		0.0088
MPJ6_DPD015[a]		Exon 7	51976695	733	GTGAATTTTGAGA/TTTGAGCTAATGA	Ile245Phe		0.0015
MPJ6_DPD016[a]		Intron 7	51976602	IVS7 + 64	CTCTACACTAAAG/TATTAACAGCAAA			0.0015
MPJ6_DPD017[a]		Exon 8	51964101	793	CTTTCAGTGAATG/AAAATGACTCTTA	Glu265Lys		0.0015
MPJ6_DPD018[a]		Intron 8	51963953	IVS8 + 91	TTCAGACATTTTC/TGTGATGAAAGTT			0.0088
MPJ6_DPD019[a]		Intron 9	51878456	IVS9–120	TTTGATAGTGACA/TCTTCATCCTGGA			0.0029
MPJ6_DPD020[b]		Exon 10	51878292	1003	ATACGGGGAGTCG/TTGATTGTACTTG	Val335Leu	*11	0.0015
MPJ6_DPD021[a]		Intron 10	51878143	IVS10 + 24	CCATCAGAAAATA/GTGGAGTTGTACT			0.0015
MPJ6_DPD022[c]		Intron 10	51858934	IVS10-15	TTTCTTCTCTGTT/CCTGTTTTGTTTT			0.018
MPJ6_DPD023[a]		Intron 12	51800901	IVS12-11	AAGTATTGGTTTG/ATATTTTGCAGTC			0.038
MPJ6_DPD024[a]		Intron 12	51800899	IVS12-9	GTATTGGTTTGTA/GTTTTGCAGTCAC			0.0073
MPJ6_DPD025[a]		Exon 13	51800872	1543	TATGGAGCTTCCG/ATTTCTGCCAAGC	Val515Ile		0.0015
MPJ6_DPD026[a]		Exon 13	51800843	1572	ACTACCCCTCTTT/GTACACTCCTATT	Phe524Leu		0.0015
MPJ6_DPD027[g]	rs1801159	Exon 13	51800788	1627	GGATTGAAGTTTA/GTAAATCCTTTTG	Ile543Val	*5	0.283
MPJ6_DPD028[a]		Exon 13	51800749	1666	ACTCCAGCCACCA/CGCACATCAATGA	Ser556Arg		0.0015
MPJ6_DPD029	rs2786783	Intron 13	51800636	IVS13 + 39	AGAAATGTCTATC/TATATATTTTAAT			0.283
MPJ6_DPD030	rs2811178	Intron 13	51800635	IVS13 + 40	GAAATGTCTATCG/ATATATTTTAATT			0.179
MPJ6_DPD031[a]		Intron 13	51735220_51735219	IVS13-47_-48	ATAAAGATTATA-/TAGCTTTTCTTTGT			0.0015
MPJ6_DPD032[a]		Exon 14	51735161	1752	GGACATTGTGACA/GAATGTTTCCCCC	Thr584Thr		0.0015
MPJ6_DPD033[f]		Exon 14	51735139	1774	CCCAGAATCATCC/TGGGGAACCACCT	Arg592Trp		0.0015
MPJ6_DPD034[g]	rs17376848	Exon 14	51735017	1896	AAAGGCTGACTTT/CCCAGACAACGTA	Phe632Phe		0.139
MPJ6_DPD035[a]		Intron 14	51734989	IVS14 + 19	GTGATTTAACATC/ATAAAACAAGAGA			0.0088
MPJ6_DPD036[a]		Intron 14	51734908	IVS14 + 100	TTAATGTGTATAT/GTTTATTAAAGAA			0.0015
MPJ6_DPD037[a]		Intron 14	51667533	IVS14-123	GATTTATTTTTTC/AACAGTTTGAAAA			0.155
MPJ6_DPD038[a]		Intron 14	51667431	IVS14-21	TGAACTTATATTC/ATTTTGTTTTTCT			0.0015
MPJ6_DPD039[d]		Intron 15	51667267	IVS15 + 75	TAAAGAGCTGCCA/GTGAGAAATAATA			0.155
MPJ6_DPD040[a]		Intron 16	51591373	IVS16-127	GGAATTTGAGAAA/GTATATCATGTAG			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*.

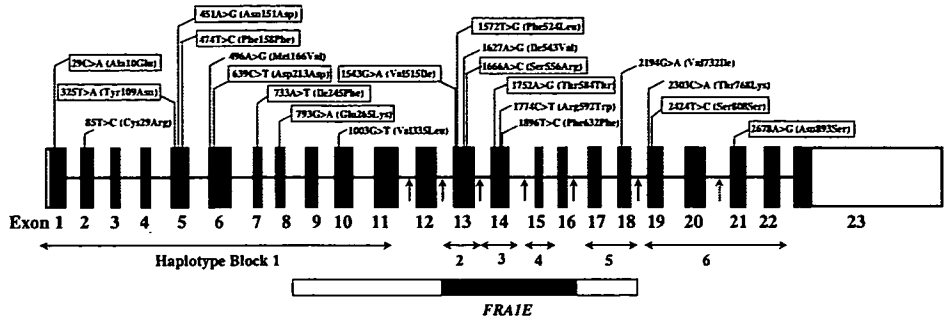

Fig. 1 Twenty-one variations detected in the coding exons are depicted in the schematic diagram of the *DPYD* gene. Fourteen novel variations are enclosed by *squares*. The recombination spots were estimated based on the LD profiles obtained from Japanese data in the HapMap project and indicated by *arrows*. 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

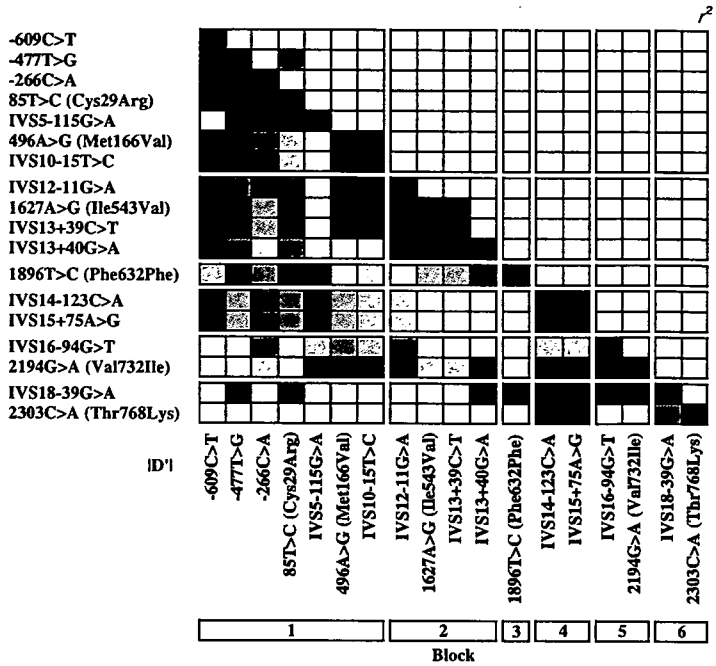

Fig. 2 Linkage disequilibrium (*LD*) analysis of *DPYD*. Pairwise LD between 18 common SNPs (>0.01 in allele frequencies) is expressed as r^2 (*upper*) and |D'| (*lower*) by a 10-graded *blue color*. The *denser color* indicates higher linkage. The haplotype block partition based on LD measure |D'| of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

Nucleotide change	-449C>T	-477T>G	-266C>A	-243G>A	29C>A	85T>C	IVS1+15T>C	IVS1+23A>G	325T>A	451A>G	471T>C	IVS3-115G>A	496A>G	639C>T	733A>T	IVS7+44G>T	793G>A	IVS8+91C>T	IVS8-128A>T	988G>T	IVS18+24A>G	IVS18-157T>C	Number	Frequency
Amino acid change					A10E	C29R			Y109N	N151D	F158F		N166V	D213D	E245F		E265K			V335I				
Haplotype *1	*1a																						553	0.818
	*1b																						31	0.045
	*1c																						14	0.021
	*1d																						13	0.019
	*1e																						6	0.0088
	*1f																						6	0.0088
	*1g																						3	0.0044
	*1h																						2	0.0029
	*1i																						2	0.0029
	*1j																						2	0.0029
	*1k																						1	0.0015
	*1l																						1	0.0015
*9	*9c																						15	0.022
	*9d																						4	0.0059
	*9e																						1	0.0015
*11	*11a																						1	0.0015
*166V	*166Va																						8	0.012
	*166Vb																						3	0.0044
*151D	*151D																						6	0.0088
*109N	*109N																						2	0.0029
*10Z	*10Z																						1	0.0015
*265K	*265K																						1	0.0015
*245F	*245F																						1	0.0015

b) Block 2

Nucleotide change			IVS12-11G>A	IVS12-9A>G	1543G>A	1572T>G	1627A>G	1666A>C	IVS13+39C>T	IVS13+40G>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									167	0.245
		*5b									26	0.038
	*515I	*515I									1	0.0015
	*524L	*524L									1	0.0015
	*556R	*556R§									1	0.0015

c) Block 3

Nucleotide change			IVS13-47_48insTA	1752A>G	1774C>T	1896T>C	IVS14+19C>A	IVS14+100T>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							1	0.0015

d) Block 4

Nucleotide change			IVS14-123C>A	IVS14-21C>A	IVS15+75A>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						10	0.015

f) Block 6

Nucleotide change			IVS18-39G>A	2303C>A	2424T>C	2678A>G	IVS21+80C>G	IVS21+113T>A	IVS21+136G>C	IVS21+162T>G	IVS22+129A>G	IVS22-69G>A	IVS22-58G>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												19	0.028
	*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

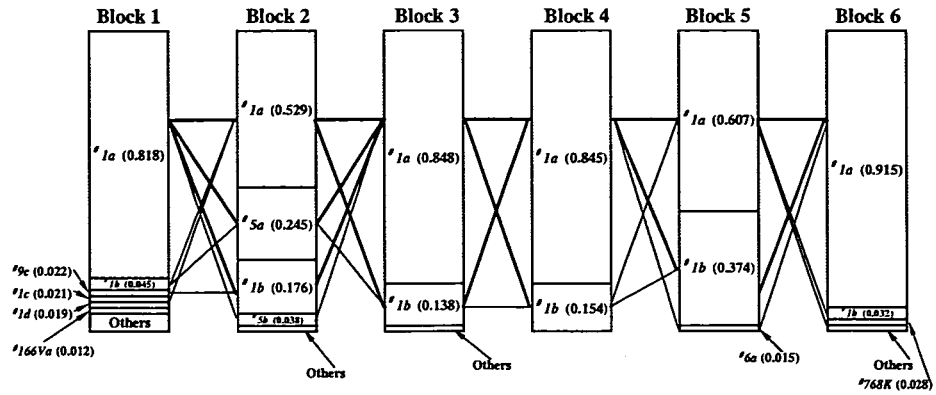

Fig. 4 The combinations of block haplotypes in Japanese. *Thick lines* represent combinations with frequencies over 10%, and *thin lines* represent combinations with frequencies of 1.0–9.9%

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