

Screening of Dihydropyrimidine Dehydrogenase Genetic Variants by Direct Sequencing in Different Ethnic Groups

Joong-Gon Shin,¹ Hyun Sub Cheong,²
Jason Yongha Kim,¹ Lyoung Hyo Kim,²
Chang Soo Han,² Ji On Kim,²
Hae Deun Kim,³ Young Hoon Kim,³
Myeon Woo Chung,³ Soon Young Han,⁴
and Hyoung Doo Shin^{1,2}

¹Department of Life Science, Sogang University, Seoul; ²Department of Genetic Epidemiology, SNP Genetics, Inc., Seoul; ³Clinical Research Division, and ⁴Toxicological Evaluation and Research Department, National Institute of Food and Drug Safety Evaluation, Osong Health Technology Administration Complex, Osong, Korea

Received: 28 December 2012
Accepted: 7 June 2013

Address for Correspondence:
Hyoung Doo Shin, PhD
Department of Life Science, Sogang University,
35 Baekbeom-ro, Mapo-gu, Seoul 121-742, Korea
Tel: +82.2-705-8615, Fax: +82.2-3273-1680
E-mail: hdshin@sogang.ac.kr

This research was supported by a grant (11182MFD5658) from Ministry of Food and Drug Safety.

Dihydropyrimidine dehydrogenase (DPYD) is an enzyme that regulates the rate-limiting step in pyrimidine metabolism, especially catabolism of fluorouracil, a chemotherapeutic agent for cancer. In order to determine the genetic distribution of *DPYD*, we directly sequenced 288 subjects from five ethnic groups (96 Koreans, 48 Japanese, 48 Han Chinese, 48 African Americans, and 48 European Americans). As a result, 56 polymorphisms were observed, including 6 core polymorphisms and 18 novel polymorphisms. Allele frequencies were nearly the same across the Asian populations, Korean, Han Chinese and Japanese, whereas several SNPs showed different genetic distributions between Asians and other ethnic populations (African American and European American). Additional *in silico* analysis was performed to predict the function of novel SNPs. One nonsynonymous SNP (+199381A > G, Asn151Asp) was predicted to change its polarity of amino acid (Asn, neutral to Asp, negative). These findings would be valuable for further research, including pharmacogenetic and drug responses studies.

Key Words: Ethnic Group; Pharmacogenetics; Dihydropyrimidine Dehydrogenase; Fluorouracil

INTRODUCTION

Pharmacogenetics focuses on identifying the role of a gene of interest that mediates drug-dependent mechanisms or triggers adverse effects. Therefore, dealing with the gene of interest is important to predict individual drug responses and toxicities. Genetic variations in genes of interest that interact with a drug may contribute to inter-individual differences in drug responses and play an important role in the designing of drugs that act on individuals with risk alleles. Recent pharmacogenetic studies have highlighted the role of genetic variations in several genes such as the *UGT* family, *CYP* family, *EGFR*, and Dihydropyrimidine dehydrogenase (*DPYD*) (1, 2).

DPYD is the initial and rate-limiting enzyme of the pyrimidine bases metabolic pathways, particularly fluorouracil (5-FU) catabolism. Recent studies have revealed that more than 80% of the medicated 5-FU, a commonly used chemotherapeutic agent for solid carcinoma, is rapidly degraded through the catabolism pathway (3, 4). Genetic variations of *DPYD* can cause an enzyme deficiency state, which results in severe toxicity or

other adverse side effects such as DNA damage or RNA damage caused by imbalance of the nucleotide pool (5-7). Several genetic variations have been reported in previous studies. *DPYD**2A (rs3918290), located in the intron site, plays the role of an alternative splicing variant, and other polymorphisms such as *DPYD**9A (rs1801265), *DPYD**7 (rs72549309), *DPYD**8 (rs1801266), *DPYD**9B (rs1801267), and *DPYD**10 (rs1801268), affect DPYD enzyme activity in other ways (8-14).

In this study, we directly sequenced the *DPYD* whole gene in 288 subjects (96 Korean, 48 Japanese, 48 Chinese, 48 African Americans, and 48 European Americans). We also analyzed the linkage disequilibrium (LD) structures and minor allele frequencies (MAFs) of the discovered single nucleotide polymorphisms (SNPs) for the gene among the different ethnicities.

MATERIALS AND METHODS

Study subjects

DNA samples from 96 unrelated Korean individuals was provided by Soonchunhyang University, Bucheon, Korea. DNA

samples from other ethnic groups was obtained from a large panel of anonymous, unrelated DNA samples from the Human Variation Panel, available through the Coriell Institute for Medical Research (Camden, NJ, USA). We specifically used sets of DNA samples obtained from four distinct ethnic groups residing in the USA, including 48 Han Chinese, 48 Japanese, 48 African Americans, and 48 European Americans individuals. The sample size was sufficient to achieve the ethnic diversity (15).

Sequencing analysis of DPYD

Promoter, all exons, and exon-intron boundaries were PCR-amplified and directly sequenced using the ABI PRISM 3730 genetic analyzer (Applied Biosystems, Foster City, CA, USA). Primers for the amplification and sequencing analysis were designed using Primer3 software (<http://frodo.wi.mit.edu>) based on the sequence of *DPYD*. The coding sequence of the gene was compared with a GenBank sequence (Ref. genome seq.: NG_008807.1). Information on the primers is listed in Supplementary Table 2. Sequence variants were verified by chromatograms using SeqMan software (DNASTAR, Madison, WI, USA).

Statistical analysis

The chi-square tests were used to determine whether individual variants were in Hardy-Weinberg equilibrium at each locus in each population. Fisher's exact test was calculated by using the Statistical Analysis System 9.2 (SAS). For in silico analysis, we used FastSNP (<http://fastsnp.ibms.sinica.edu.tw>), Expasy (<http://expasy.org/tools>), and UTRScan (<http://itbtools.ba.itb.cnr.it/utrscan>) programs to predict the function of novel SNPs.

Ethics statement

The protocol and consent forms of this study were reviewed and approved by the institutional review board of Sogang University (2010_690). Informed consent was submitted by the subjects.

RESULTS

In order to discover *DPYD* SNPs, we directly sequenced 288 samples from five ethnic groups (Korean, Han Chinese, Japanese, African American and European American). As a result, 56 SNPs were found, including 18 novel SNPs. Among the novel SNPs, five (+199381A > G, Asn151Asp; +199404T > C, Phe158Phe; +221378A > G, Val162Val; +221531C > T, Asp213Asp; and +841847T > C, Leu993Arg) were located in coding regions (Table 1).

MAFs and relative physical coordinates of all SNPs are shown in Table 1 and Supplementary Fig. 1. Allele frequencies were nearly the same among the Korean, Han Chinese, and Japanese samples, whereas several SNPs showed different genetic distributions between Asians and other ethnic populations (African

American and European American). Among those SNPs, the frequency of a core marker, *9A (rs1801265), in Asian populations was somewhat lower than in the African American and European American samples (MAF: Korean = 0.016, Han Chinese = 0.043, Japanese = 0.065, African American = 0.490, European American = 0.177). In contrast, other core markers, *7 (rs72549309), *8 (rs1801266), *2A (rs3918290), *9B (rs1801267), and *10 (rs1801268), were monomorphic in all the studied populations.

In order to find significant differences in allele frequencies between Korean and other ethnic groups, Fisher's exact test was additionally conducted (Supplementary Fig. 1). The test results indicated that there were significant differences between Asians and other ethnic populations (African American and European American) in the six SNPs (*9A, rs668296, rs2811178, rs56160474, rs291592, and rs291593). Among them, a core marker *9A (rs1801265) showed the most significant differences ($P = 6.61 \times 10^{-19}$ and 2.47×10^{-6} for Korean vs African American and Korean vs European American samples, respectively). Moreover, the reversal of major and minor alleles was observed in rs291592 (C allele is major in Asians, but minor in African American and European American). Also, genetic difference was also observed within the Asians in rs291593 (T allele is common in Korean and Han Chinese, whereas it is rarely found in Japanese, African American, and European American). Detailed information about core markers such as star allele nomenclature, position, allele change, amino acid change, and any known roles in pharmacogenetics is presented in Table 1.

DISCUSSION

DPYD is an enzyme that takes part in a rate-limiting step of 5-FU catabolism. Previous studies have shown that the enzyme deficiency state of the 5-FU degradation pathway causes damage and degeneration of the central nervous system (8, 14, 16). Thus *DPYD* is known as a biomarker of severe toxicity in chemotherapeutic agents. Several *DPYD* polymorphisms have been reported as clinical loci associated with a reduced level of enzyme activity and severe 5-FU toxicity, and these polymorphisms are called "core markers". The most studied core markers are *DPYD**9A (rs1801265) and *DPYD**2A (rs3918290) (17-19).

The core marker *9A (rs1801265) is located in the coding region and induces amino acid change (cysteine to arginine) that may affect enzyme activity. It is relatively common in Caucasian populations (MAF > 10%), although *DPYD* enzyme activity is not affected by the polymorphism (20-23). This polymorphism is rare in Asian populations, but previous studies have reported that the incidence of clinical presentation of enzyme deficiency caused by heterozygous *9A (rs1801265) is significantly higher than the wild type ($P < 0.05$) in the Chinese population (23-26). In this study, frequency of *9A (rs1801265) showed a similar trend

Table 1. Allele frequency of *DPYD* in study (n=288)

SNPID		Coordinate	Allele [†]	Activity of <i>DPYD</i> (14)	Position	AA Change	Alleles	Minor allele frequency [‡]				
Novel SNP	rs #							KR	HC	JP	AA	EA
	rs117514263	98387877			Promoter	.	G > A	-	-	0.010	-	-
-609C > T		98387224			Promoter	.	C > T	0.068	0.043	0.010	-	-
-442A > C		98387057			Promoter	.	A > C	-	0.011	-	-	-
	rs72981745	98387035			Promoter	.	C > T	-	-	-	0.052	-
	rs61787828	98386955			Promoter	.	T > G	0.005	0.011	0.062	0.010	0.125
	rs1801265	98348885	9A*	Normal	Exon2	C29R	A > G	0.016	0.043	0.065	0.490	0.177
+180351C > T		98206264			Intron3	.	C > T	-	-	-	0.010	-
+180409T > C		98206206			Intron3	.	T > C	-	0.010	-	-	-
+180600A > G		98206015			Intron4	.	A > G	-	0.010	-	-	-
	rs72549309	98205967	7*	Uncertain	Exon4	I101I	TCAT_insdel	-	-	-	-	-
+199381A > G		98187234			Exon5	N151D	A > G	0.005	-	-	-	-
+199404T > C		98187211			Exon5	F158F	T > C	0.005	0.021	-	-	-
+199994T > A		98186621			Intron5	.	T > A	-	0.021	0.031	0.010	-
+200113A > G		98186502			Intron5	.	A > G	-	-	-	0.031	-
+200207C > T		98186408			Intron5	.	C > T	0.005	-	-	-	-
	rs56066952	98186232			Intron5	.	A > G	-	-	-	-	0.010
	rs55684412	98186229			Intron5	.	A > G	0.026	0.021	-	0.156	0.115
+201152G > A		98185463			Intron5	.	G > A	0.010	-	-	-	-
+221378A > G		98165237			Exon7	V162V	A > G	0.005	-	-	-	-
	rs2297595	98165091			Exon7	M166V	A > G	0.031	0.021	0.021	0.031	0.094
+221531C > T		98165084			Exon7	D213D	C > T	-	-	0.042	-	-
	rs115232898	98165030			Exon7	Y186C	A > G	-	-	-	0.031	-
	rs6668296	98164768			Intron7	.	G > A	0.031	0.022	0.021	0.219	0.117
	rs1801266	98157332	8*	Decrease	Exon8	R235W	C > T	-	-	-	-	-
	rs2786491	98060753			Intron8	.	C > T	-	-	-	0.094	-
	rs114968502	98060579			Intron9	.	A > G	-	-	-	0.010	-
	rs56293913	98039541			Intron10	.	T > C	0.021	0.021	0.021	0.043	0.146
	rs61622928	98039437			Exon11	M406I	G > A	-	-	-	0.064	-
	rs57918000	98015269			Exon12	N457N	C > T	-	-	-	0.022	-
	rs116364703	98015146			Exon12	Q498Q	A > G	-	-	-	0.011	-
	rs55699321	97981508			Intron12	.	G > A	0.043	0.042	0.021	-	-
	rs56056384	97981506			Intron12	.	A > G	-	0.010	0.011	-	-
	rs1801158	97981421	4*	Normal	Exon13	N534S	G > A	-	-	-	-	0.096
	rs1801159	97981395	5*	Normal	Exon13	I543V	A > G	0.297	0.292	0.219	0.200	0.188
	rs2786783	97981243			Intron13	.	C > T	0.287	0.281	0.219	0.138	0.177
	rs2811178	97981242			Intron13	.	G > A	0.165	0.281	0.250	0.475	0.438
	rs3918290	97915614	2A*	Decrease	Intron14	Splicing variant	C > T	-	-	-	-	-
	rs12078940	97848041			Intron14	.	A > C	-	-	-	0.073	-
+538604C > T		97848011			Intron15	.	C > T	-	-	-	0.010	-
	rs72728438	97847874			Intron15	.	A > G	0.200	0.198	0.250	0.229	0.260
	rs74104343	97770937			Intron17	.	T > C	-	-	-	0.022	-
	rs1801160	97770920	6*	Normal	Exon18	V732I	G > A	0.016	-	0.022	0.022	0.083
	rs60511679	97770919			Exon18	V732G	T > G	-	-	-	0.011	-
	rs12137711	97700589			Intron18	.	G > A	0.047	0.011	0.033	0.021	0.074
	rs1801267	97564154	9B*	Decrease	Exon20	R886H	G > A	-	-	-	-	-
+838658C > T		97547957			Intron21	.	C > T	0.010	-	-	-	-
+841847T > G		97544768			Exon22	L993R	T > G	-	-	-	0.010	-
	rs1801268	97544627	10*	Uncertain	Exon22	V995F	G > T	-	-	-	-	-
	rs114096998	97544543			Exon22	P1023T	C > A	-	-	-	0.052	-
	rs56160474	97544258			3'UTR	.	T > C	-	-	-	0.115	0.177
+842533C > T		97544082			3'UTR	.	C > T	-	-	-	-	0.010
	rs1042482	97543959			3'UTR	.	G > A	0.174	0.104	0.156	0.010	0.094
	rs291592	97543764			3'UTR	.	G > G	0.005	0.031	0.042	0.323	0.417
	rs291593	97543752			3'UTR	.	C > T	0.395	0.375	0.479	0.260	0.271
	rs41285690	97543470			3'UTR	.	A > G	-	-	-	-	0.021
+843187C > A		97543428			3'down	.	C > A	-	-	-	-	0.010

*Alleles of core markers were verified from previous studies. (14, 19). -, monomorphic; †, core SNP; ‡, major and minor alleles determined by frequency of all subjects. KR, Korean; HC, Han Chinese; JP, Japanese; AA, African American; EA, European American.

to that of previous studies (Supplementary Table 1).

Another important clinical locus, *2A (rs3918290), is in the splicing recognition sequence of the intron region. Genetic variation on this site can lead to deletion of the 165 base pair corresponding to the nearby coding region (exon 14), and consequently, the ceasing of enzyme activity (14). However, enzyme deficiency caused by *2A (rs3918290) is known to be rare in both Caucasians and Asians (Supplementary Table 1) (22, 24, 27, 28).

In addition, direct sequencing of *DPYD* in our study led to the discovery of a number of novel SNPs. In order to predict the function of the novel SNPs, in silico analyses were conducted according to the position of the polymorphisms. As a result, +842533C > T, located in the 3'-untranslated region (3'UTR), was predicted to introduce an internal ribosome entry site (IRES) motif. IRES allows the ribosome to bind to the mRNA internally and translate it rather than binding to the 5' cap as normally occurs (29-31). Although IRES is involved in 5'UTR mediated translation initiation, conserved secondary structures of IRES can influence the 3'UTR stability (32, 33). Moreover, a recent study in the Japanese population discovered a number of 3'UTR novel SNPs which were located near microRNA target sites (34). In addition, a previous study reported that binding of microRNA to 3'UTR negatively regulates the mRNA of target gene (35). Moreover, two nonsynonymous SNPs (+199381A > G, Asn151Asp and +841847T > G, Leu993Arg) were also found in our study. Polarity of the amino acid was affected by the charge of its side-chain (36). Polarity alteration between amino acids was predicted that Asn151Asp may affect protein structure or function (Asn, neutral; Asp, negative). Although frequency of Asn151Asp was low or monomorphic in our study subjects, it can be associated with the enzyme deficiency state.

Although PCR direct sequencing method was adopted for estimating the frequency differences across different ethnic groups and 48 samples per population is large enough to discover novel SNPs, a larger sample would be required to achieve more detailed screening result. Also, no functional study was conducted to further confirm the SNPs' role, although in silico analyses were performed to compensate for the lack of functional analysis to an extent.

In summary, the present study analyzed *DPYD* by directly sequencing 288 subjects from five ethnic groups. This yielded 56 SNPs, including 9 core SNPs and 18 novel SNPs. Moreover, we predicted the function of novel SNPs using in silico analyses. Although a lack of functional studies might be a limitation of the study, the results could make a valuable contribution to further research, especially pharmacogenetic studies of drug responses.

DISCLOSURE

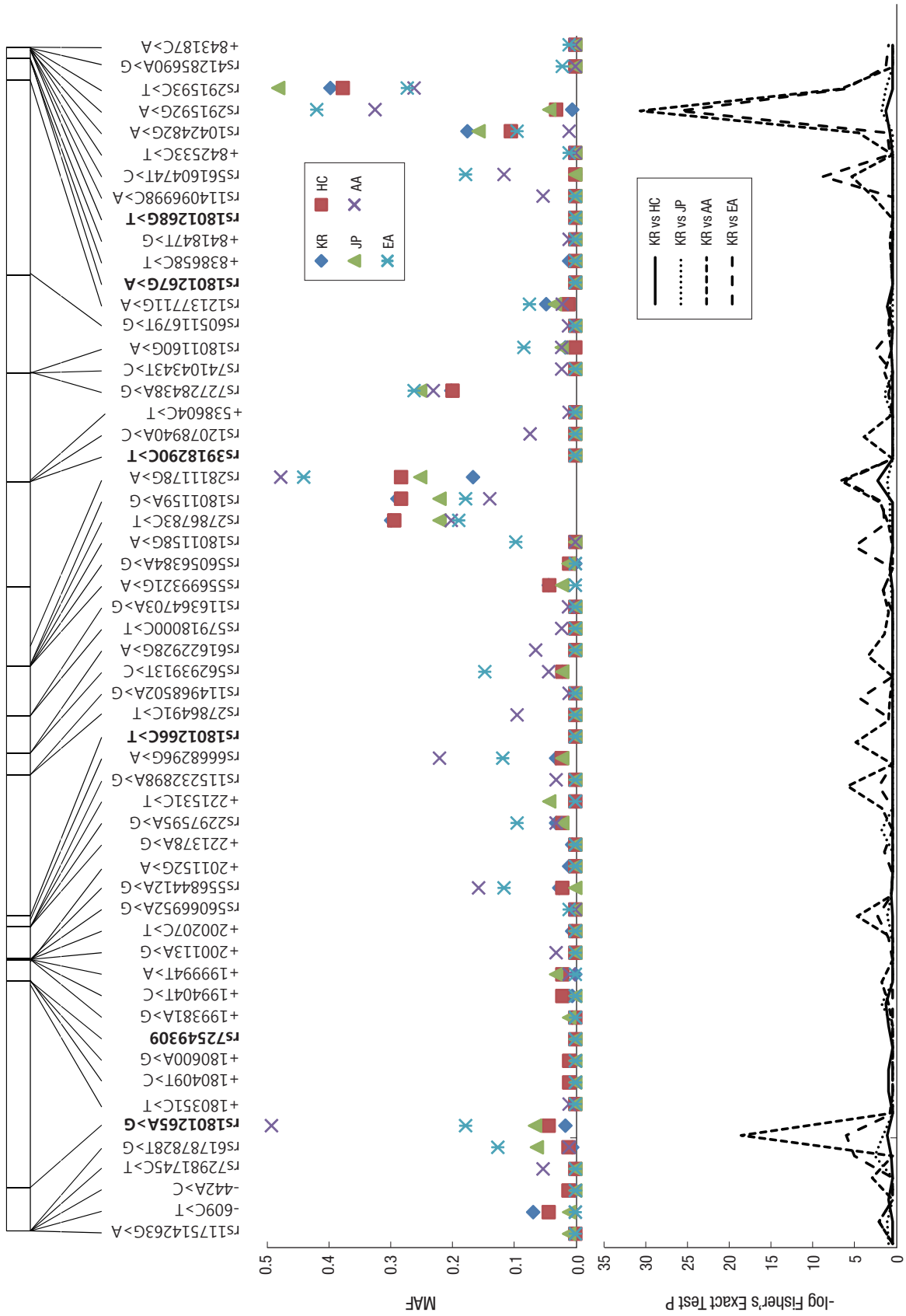
The authors have no conflicts of interest to disclose.

REFERENCES

1. Savonarola A, Palmirotta R, Guadagni F, Silvestris F. *Pharmacogenetics and pharmacogenomics: role of mutational analysis in anti-cancer targeted therapy. Pharmacogenomics J* 2012; 12: 277-86.
2. Kristyanto H, Utomo AR. *Pharmacogenetic application in personalized cancer treatment. Acta Med Indones* 2010; 42: 109-15.
3. Diasio RB, Harris BE. *Clinical pharmacology of 5-fluorouracil. Clin Pharmacokinet* 1989; 16: 215-37.
4. Milano G, McLeod HL. *Can dihydropyrimidine dehydrogenase impact 5-fluorouracil-based treatment? Eur J Cancer* 2000; 36: 37-42.
5. Amstutz U, Froehlich TK, Largiadèr CR. *Dihydropyrimidine dehydrogenase gene as a major predictor of severe 5-fluorouracil toxicity. Pharmacogenomics* 2011; 12: 1321-36.
6. Longley DB, Harkin DP, Johnston PG. *5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer* 2003; 3: 330-8.
7. Mattison LK, Soong R, Diasio RB. *Implications of dihydropyrimidine dehydrogenase on 5-fluorouracil pharmacogenetics and pharmacogenomics. Pharmacogenomics* 2002; 3: 485-92.
8. Cerić T, Obralić N, Kapur-Pojškić L, Macić D, Beslija S, Pasić A, Cerić S. *Investigation of IVS14 + 1G > A polymorphism of DPYD gene in a group of Bosnian patients treated with 5-Fluorouracil and capecitabine. Bosn J Basic Med Sci* 2010; 10: 133-9.
9. Ezzeldin HH, Lee AM, Mattison LK, Diasio RB. *Methylation of the DPYD promoter: an alternative mechanism for dihydropyrimidine dehydrogenase deficiency in cancer patients. Clin Cancer Res* 2005; 11: 8699-705.
10. Johnson MR, Wang K, Diasio RB. *Profound dihydropyrimidine dehydrogenase deficiency resulting from a novel compound heterozygote genotype. Clin Cancer Res* 2002; 8: 768-74.
11. Van Kuilenburg AB, Dobritzsch D, Meinsma R, Haasjes J, Waterham HR, Nowaczyk MJ, Maropoulos GD, Hein G, Kalhoff H, Kirk JM, et al. *Novel disease-causing mutations in the dihydropyrimidine dehydrogenase gene interpreted by analysis of the three-dimensional protein structure. Biochem J* 2002; 364: 157-63.
12. Van Kuilenburg AB, Vreken P, Beex LV, Meinsma R, Van Lenthe H, De Abreu RA, van Gennip AH. *Heterozygosity for a point mutation in an invariant splice donor site of dihydropyrimidine dehydrogenase and severe 5-fluorouracil related toxicity. Eur J Cancer* 1997; 33: 2258-64.
13. Ezzeldin H, Johnson MR, Okamoto Y, Diasio R. *Denaturing high performance liquid chromatography analysis of the DPYD gene in patients with lethal 5-fluorouracil toxicity. Clin Cancer Res* 2003; 9: 3021-8.
14. Ezzeldin H, Diasio R. *Dihydropyrimidine dehydrogenase deficiency, a pharmacogenetic syndrome associated with potentially life-threatening toxicity following 5-fluorouracil administration. Clin Colorectal Cancer* 2004; 4: 181-9.
15. Thompson D, Stram D, Goldgar D, Witte JS. *Haplotype tagging single nucleotide polymorphisms and association studies. Hum Hered* 2003; 56: 48-55.
16. Thorn CF, Marsh S, Carrillo MW, McLeod HL, Klein TE, Altman RB. *PharmGKB summary: fluoropyrimidine pathways. Pharmacogenet Genomics* 2011; 21: 237-42.
17. McLeod HL, Collie-Duguid ES, Vreken P, Johnson MR, Wei X, Sapone A, Diasio RB, Fernandez-Salguero P, van Kuilenburg AB, van Gennip AH, et al. *Nomenclature for human DPYD alleles. Pharmacogenetics* 1998; 8:

- 455-9.
18. Saif MW, Ezzeldin H, Vance K, Sellers S, Diasio RB. *DPYD*2A mutation: the most common mutation associated with DPD deficiency. Cancer Chemother Pharmacol* 2007; 60: 503-7.
 19. Van Kuilenburg AB, Vreken P, Abeling NG, Bakker HD, Meinsma R, Van Lenthe H, De Abreu RA, Smeitink JA, Kayserili H, Apak MY, et al. *Genotype and phenotype in patients with dihydropyrimidine dehydrogenase deficiency. Hum Genet* 1999; 104: 1-9.
 20. Vreken P, Van Kuilenburg AB, Meinsma R, van Gennip AH. *Dihydropyrimidine dehydrogenase (DPD) deficiency: identification and expression of missense mutations C29R, R886H and R235W. Hum Genet* 1997; 101: 333-8.
 21. Kelemen LE, Goodman MT, McGuire V, Rossing MA, Webb PM, Köbel M, Anton-Culver H, Beesley J, Berchuck A, Brar S, et al. *Genetic variation in TYMS in the one-carbon transfer pathway is associated with ovarian carcinoma types in the Ovarian Cancer Association Consortium. Cancer Epidemiol Biomarkers Prev* 2010; 19: 1822-30.
 22. Maekawa K, Saeki M, Saito Y, Ozawa S, Kurose K, Kaniwa N, Kawamoto M, Kamatani N, Kato K, Hamaguchi T, et al. *Genetic variations and haplotype structures of the DPYD gene encoding dihydropyrimidine dehydrogenase in Japanese and their ethnic differences. J Hum Genet* 2007; 52: 804-19.
 23. Collie-Duguid ES, Etienne MC, Milano G, McLeod HL. *Known variant DPYD alleles do not explain DPD deficiency in cancer patients. Pharmacogenetics* 2000; 10: 217-23.
 24. Cho HJ, Park YS, Kang WK, Kim JW, Lee SY. *Thymidylate synthase (TYMS) and dihydropyrimidine dehydrogenase (DPYD) polymorphisms in the Korean population for prediction of 5-fluorouracil-associated toxicity. Ther Drug Monit* 2007; 29: 190-6.
 25. Zhang H, Li YM, Zhang H, Jin X. *DPYD*5 gene mutation contributes to the reduced DPYD enzyme activity and chemotherapeutic toxicity of 5-FU: results from genotyping study on 75 gastric carcinoma and colon carcinoma patients. Med Oncol* 2007; 24: 251-8.
 26. Yamaguchi K, Arai Y, Kanda Y, Akagi K. *Germline mutation of dihydropyrimidine dehydrogenase gene among a Japanese population in relation to toxicity to 5-Fluorouracil. Jpn J Cancer Res* 2001; 92: 337-42.
 27. Amstutz U, Farese S, Aebi S, Largiadèr CR. *Dihydropyrimidine dehydrogenase gene variation and severe 5-fluorouracil toxicity: a haplotype assessment. Pharmacogenomics* 2009; 10: 931-44.
 28. Savva-Bordalo J, Ramalho-Carvalho J, Pinheiro M, Costa VL, Rodrigues A, Dias PC, Veiga I, Machado M, Teixeira MR, Henrique R, et al. *Promoter methylation and large intragenic rearrangements of DPYD are not implicated in severe toxicity to 5-fluorouracil-based chemotherapy in gastrointestinal cancer patients. BMC Cancer* 2010; 10: 470.
 29. Yaman I, Fernandez J, Liu H, Caprara M, Komar AA, Koromilas AE, Zhou L, Snider MD, Scheuner D, Kaufman RJ, et al. *The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. Cell* 2003; 113: 519-31.
 30. Hellen CU, Sarnow P. *Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev* 2001; 15: 1593-612.
 31. Stoneley M, Willis AE. *Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. Oncogene* 2004; 23: 3200-7.
 32. Chatterjee S, Pal JK. *Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. Biol Cell* 2009; 101: 251-62.
 33. Conklin D, Jonassen I, Aasland R, Taylor WR. *Association of nucleotide patterns with gene function classes: application to human 3' untranslated sequences. Bioinformatics* 2002; 18: 182-9.
 34. Hirota T, Date Y, Nishibatake Y, Takane H, Fukuoka Y, Taniguchi Y, Burioka N, Shimizu E, Nakamura H, Otsubo K, et al. *Dihydropyrimidine dehydrogenase (DPD) expression is negatively regulated by certain microRNAs in human lung tissues. Lung Cancer* 2012; 77: 16-23.
 35. Ambros V. *The functions of animal microRNAs. Nature* 2004; 431: 350-5.
 36. Weinshilboum R, Wang L. *Pharmacogenomics: bench to bedside. Nat Rev Drug Discov* 2004; 3: 739-48.

Supplementary Fig. 1. Minor allele frequency of *DPYD* among Korean, Han Chinese, Japanese, African American and European American samples (upper panel). Fisher's exact test of allele distribution between Korean and other ethnic groups (lower panel). Relative physical coordinates of SNPs are shown at top of the figure. Bolded SNPs indicate pharmacogenetic core SNPs. Novel SNPs are labeled with their location and allele change. KR, Korean; HC, Han Chinese; JP, Japanese; AA, African American; EA, European American.



Supplementary Table 1. Comparison of minor allele frequency of *9A (rs1801265) and *2A (rs3918290) between present study and previous studies

		MAF	
		*9A (rs1801265)	*2A (rs3918290)
Present study	KR	0.016	-
	HC	0.043	-
	JP	0.065	-
	AA	0.49	-
	EA	0.177	-
Previous studies	KR	0.025 (24)	0 (24)
	HC	0.05 (25)	NR
	JP	0.037 (26)	0 (23)
	CA	0.16 (19), 0.21 (21), 0.34 (23)	0.013, 0.007 (23)

MAF, minor allele frequency; -, monomorphic; NR, not reported; KR, Korean; HC, Han Chinese; JP, Japanese; AA, African American; EA, European American; CA, Caucasian.

Supplementary Table 2. Primer information of *DPYD*

Name	Forward seq. (5'-3')	Reverse seq. (5'-3')
DPYD_01P	TGCCCAGTGTATAAACGAATGT	TCCATACCAGATTGAGCTTTAGAA
DPYD_02P	GAGCAGGTTATTACGCGGT	GACACTGTGCTAGGTTCTCCG
DPYD_03P	CAAGCAGGCATCACATTTTC	GAGCAAAAACATCCACAGCC
DPYD_04P	TCTCTCCCTTCTCCTCAGCA	GGGGACTGGCCTTTAATCTC
DPYD_05P	GTGCAAGTTGTTGCTTGGAG	GAAACTTTCCCGCGTCTCTC
DPYD_06P	GAGATTAAGGCCAGTCCCC	GAAACTTTCCCGCGTCTCTC
DPYD_07P	TCAATTTAATGAATGCTACCCAA	CCAAAGTACAGCCTCAAGGG
DPYD_08P	CCAAGAGGAGTGCCAAAGAT	TGGATTTGCTAAGACAAGCTG
DPYD_09P	TGCATTGAATATTCCTATGTGTCA	TCCTTTTTGTAGCTAAGCTGCTG
DPYD_10P	TGAACTTATAGATTGGAAGCAGGTC	GCAGGAAGGGAAATAACCCTG
DPYD_11P	TGCCCGTCAACCTAGATTTT	CACACACCTCTCATGACCAGA
DPYD_12P	TCACAGATGCAGTAAGTAAAGG	GTTCCACGAGGACATTTT
DPYD_13P	ACCACTATTCCAGAGCAG	TTTGAAGAATAACATTCGCA
DPYD_14P	CTGAAAATGACTGCTATTGCT	GCTTCAAGCCAAGTCAAAAT
DPYD_15P	TCCTCATGCATATCTTGTGTGAT	CAAGGGGAAGCATTTTCTG
DPYD_16P	GGCTGACTTTTCATTCTTTTGA	GGTGATGATGGTGTGGTGAA
DPYD_17P	TTGATTTGCTTACAGATGTTTTCC	GGTCTTAGGCAAGGTTGGGT
DPYD_18P	TGGGAATAAACTGTCTTCAATG	TGACAATTTCAACATTCTAGCGA
DPYD_19P	CTCCAATTCAGACTCCCTGA	TGAAAAACAATTCCTGAAAAGC
DPYD_20P	TTGTTTGAATCCCTGGAACA	CCTGGCCCAATTTTAAATCA
DPYD_21P	TGAATTCAGTTCAGTCTCACT	TGGGACAGAAAGGAAGGAAA
DPYD_22P	TGCAAAAATGTGAGAAGGGA	TTCCATCTTACATGGCACCC
DPYD_23P	TAATTCAAAAGCCCAAATG	TGTGTGAAATCCAAGGGACC
DPYD_24P	TTGTTGTGGTGGTGGTTGTT	GGGGCCTGTATGTCACACT
DPYD_25P	TTTGTCTTGACGCTCCAG	TGCTAGCATGAGTCCAGGTG
DPYD_26P	TGAGAAAGTAAAGTTGTGGTAATAAA	TTGAGCTTTCATATTCATCCG
DPYD_27P	GTCCAGTGACGCTGTCATCA	TAAGCCCTCAACAGGACAGG
DPYD_28P	CCAGACGGCTACTGATCCAT	AAATCACATCCAGGAGGCAC
DPYD_29P	CGGAACCTGATACCGAGAAG	GCATGCTTGCCAGTGTTCTA
DPYD_30P	AAAACAGGAAAATGCTGAGTGA	TGTTGTGGCTGATGAAATGG
DPYD_31P	TTGTCTCATAGTGTGGCTCCTC	AACTGGCAGTGAACATCCAA
DPYD_32P	TCAGCTTTTCCAAATTAACAAA	TCACGCATGAGTAGTGAATGG
DPYD_33P	TGCTATGGCAGGATAAGAAAAGC	TGAAGATGCCATGAAGAGGA
DPYD_34P	AAGTGTGCCAAATCAGTTTGA	GGGTAAGAAGGACAGGAAGTGA