J Hum Genet (2000) 45:299–302

ORIGINAL ARTICLE

Toyokazu Seki · Toshihiro Tanaka · Yusuke Nakamura

Genomic structure and multiple single-nucleotide polymorphisms (SNPs) of the thiopurine S-methyltransferase (TPMT) gene

Received: June 26, 2000 / Accepted: July 31, 2000

Abstract Thiopurine S-methyltransferase (TPMT) catalyzes the S-methylation of drugs such as azathiopurine, 6-mercaptopurine, and 6-thioguanine, which are widely prescribed for immunosuppressive or cytotoxic applications. We report here the entire genomic structure of the *TPMT* gene and the presence of 30 single-nucleotide polymorphisms (SNPs) within that structure. The gene spans a genomic region about 27 kb long and consists of nine exons. By screening its entire genomic sequence for SNPs in 48 Japanese chromosomes by direct DNA sequencing, we detected 1 SNP in the 870-bp promoter region, 26 SNPs in introns, and 3 SNPs in the 3' untranslated region (3'UTR) for investigating correlations between *TPMT* genotypes and the side-effects caused by thiopurine drugs.

Key words Single nucleotide polymorphism (SNP) · Thiopurine S-methyltransferase (TPMT) · Side effect · Genomic structure · Drug metabolism

Introduction

Thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme that catalyzes the S-methylation of aromatic and heterocyclic sulfhydryl compounds, among them thiopurine derivatives, such as 6-mercaptopurine, thioguanine, and azathiopurine. Thioguanine and 6-mercaptopurine are among the most widely used agents for treating leukemias, and azathiopurine is widely prescribed for immunosuppression (Weinshilboum and Sladek 1980; McLeod et al. 1995; Krynetski et al. 1996). Most of the nucleotide metabolites of these drugs can be methylated by TPMT to yield inactive S-methylated compounds (Elion 1989; Krynetski et al. 1995a). However, TPMT-deficient patients treated with conventional doses of thiopurine drugs are at a high risk of severe, and occasionally fatal, hematopoietic toxicity (Krynetski and Evans 1998).

In approximately 1 in 300 subjects (0.3%), TPMT activity, as measured in erythrocytes, is very low or undetectable; 6%–11% of subjects show intermediate activity, and 89%–94% show high activity (Weinshilboum and Sladek 1980; Krynetski et al. 1996; McLeod et al. 1994). Recent studies have established that differences in the activity of this enzyme reflect variations in the *TPMT* gene itself (Krynetski et al. 1995b), which has been reported to consist of ten exons and to be located on chromosome 6p22.3 (Szumlanski et al. 1996). Several variant alleles containing base substitutions have been reported to date, among them five major alleles (TPMT*1, TPMT*2, TPMT*3A, TPMT*3B, and TPMT*3C) (Tai et al. 1996; Otterness et al. 1997; Otterness et al. 1998; Ameyaw et al. 1999; Hon et al. 1999). TPMT*1 corresponds to the wild type, which encodes a highly active product. In earlier studies, the four other alleles accounted for about 80% of white or black subjects whose TPMT activity was low or intermediate (Tai et al. 1996; Otterness et al. 1997; Otterness et al. 1998; Ameyaw et al. 1999; Hon et al. 1999); the TPMT*2 allele includes a G-to-C substitution at nucleotide 238 of the cDNA. TPMT*3 alleles have been classified into three subgroups: TPMT*3A contains two substitutions (G460A and A719G); TPMT*3B has G460A alone; and TPMT*3C contains only the A719G substitution. Individuals with mutations on both alleles have little or undetectable TPMT activity, and those carrying one wild-type allele show intermediate TPMT activity. Because of the significant correlation between genetic variation and phenotype (enzyme activity), TPMT is thought to be an important determinant of toxicity associated with thiopurine medications.

However, in Asian populations, these known genetic variants are less frequent than in white or black subjects (Collie-Duguid et al. 1999). Furthermore, about a fivefold difference in TPMT activity occurs in both white and black populations, who carry two copies of the wild-type *TPMT* allele, indicating that other genetic differences may influence *TPMT* activity. We considered that single nucleotide

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

polymorphisms (SNPs) in non-coding regions might be involved in this additional control mechanism.

In the study reported here, we determined the entire genomic sequence of the *TPMT* gene and then searched for SNPs in its promoter region, and in all introns.

Materials and methods

Amplification of samples. Total genomic DNAs were isolated from peripheral leukocytes of 24 unrelated Japanese individuals by phenol/chloroform extraction. We designed polymerase chain reaction (PCR) primers to amplify each intron of the *TPMT* gene in a Gene Amp PCR system 9600 (Perkin Elmer, Foster City, CA, USA), using, in each experiment, 100 ng of mixed DNA derived from three individuals in equal amounts. Initial denaturation was at 94°C for 5 min, followed by 35 cycles of amplification at 98°C for 10 s, annealing at 60°C–64°C for 1 min, and extension at 72°C for 1–10 min. Annealing temperatures and extension times were adjusted according to the melting temperature of each pair of primers and the length of each intron.

Direct sequencing. Products obtained from the above PCR experiments were used as templates for direct sequencing and detection of SNPs, by the fluorescent dye-terminator cycle sequencing method (ABIPRISM Big Dye Terminator Cycle Sequencing Ready Reaction kit; Perkin Elmer). The forward and reverse primers were designed based on nucleotide sequences determined with a capillary-based ABI 3700 autosequencer (Perkin Elmer).

Detection of polymorphism and estimation of allele frequency. We invoked the PolyPhred computer program (Nickerson et al. 1997), which automatically detects the presence of heterozygous SNPs by fluorescence-based sequencing of PCR products. Each SNP candidate was inspected manually before being confirmed. Allele frequencies were estimated roughly by comparing the peak levels of signals.

Results

Nucleotide sequences of introns in the TPMT gene, and SNPs within the genomic sequence. Using partial information about the genomic structure of the *TPMT* gene that had been reported previously (Szumlanski et al. 1996), we amplified each presumptive intron by the PCR and determined the gene's entire genomic sequence. Although other investigators had indicated the presence of nine introns (ten exons), we were unable to find their "intron 2" (Szumlanski et al. 1996). We determined that the nine-exon *TPMT* gene spans a genomic region about 27 kb long (GenBank accession number, AB045146) (Fig. 1). We designed PCR primers to amplify DNA fragments containing no repetitive elements, by invoking the REPEAT MASKER computer program (http://ftp.genome.washington.edu/cgi.bin/RepeatMasker). Sequencing of DNA regions, and in the promoter region, detected the 30 SNPs listed in Table 1. The average frequency of SNPs detected in introns was 26 in 23,594 bp (1 in every 907 bp). The allelic frequencies listed in Table 1 were calculated from the DNA sequence data.

Table 1. Summary of 30 non-coding SNPs detected in the *TPMT* gene

Location	Position[a]	SNP	Allelic frequency
Promoter	−315	G/A	0.98:0.02
Intron 1	24	G/A	0.98:0.02
Intron 1	1934	A/T	0.52:0.48
Intron 1	2842	C/T	0.79:0.21
Intron 1	4720	G/C	0.56:0.44
Intron 1	5566	A/C	0.60:0.40
Intron 1	5708	G/A	0.73:0.27
Intron 2	709	T/A	0.98:0.02
Intron 3	35	C/T	0.67:0.33
Intron 3	274	C/T	0.79:0.21
Intron 3	868	G/T	0.55:0.45
Intron 3	1381	G/A	0.77:0.23
Intron 3	2643	C/T	0.52:0.48
Intron 3	2682	T/A	0.54:0.46
Intron 3	3095	T/A	0.65:0.35
Intron 3	3625	T/A	0.58:0.42
Intron 4	48	T/C	0.56:0.44
Intron 4	858	A/T	0.75:0.25
Intron 4	2449	G/A	0.90:0.10
Intron 4	2520	A/C	0.94:0.06
Intron 4	3580	G/A	0.83:0.17
Intron 6	642	C/T	0.65:0.35
Intron 6	954	G/A	0.63:0.37
Intron 6	3204	C/A	0.71:0.29
Intron 7	782	T/C	0.75:0.25
Intron 7	1216	A/G	0.63:0.37
Intron 8	522	T/A	0.52:0.48
Exon 9	2001	C/A	0.79:0.21
Exon 9	2352	T/A	0.65:0.35
Exon 9	2452	T/C	0.98:0.02

SNP, Single nucleotide polymorphism

[a] Nucleotide positions are counted from the first intronic nucleotide at the exon/intron junction (for SNPs in the promoter region or in exon 9, the numbers indicate nucleotide positions with respect to the translation-initiation site)

Discussion

Thiopurine S-methyltransferase (TPMT) catalyzes the S-methylation of aromatic and heterocyclic sulfhydryl compounds, such as 6-mercaptopurine, thioguanine, and azathiopurine, three drugs that are commonly prescribed for antitumor therapy or immunosuppressive treatment. Most of the nucleotide metabolites of these drugs can be methylated and inactivated by TPMT (Elion 1989; Krynetski et al. 1995a). Several variant alleles of the *TPMT* gene that affect its catalytic activity have been reported (Tai et al. 1996; Otterness et al. 1997; Otterness et al. 1998; Ameyaw et al. 1999; Hon et al. 1999). Although the proportion of Asian individuals showing low TPMT activity is similar to that in European or American populations, the frequencies of the known genetic variants are very low in

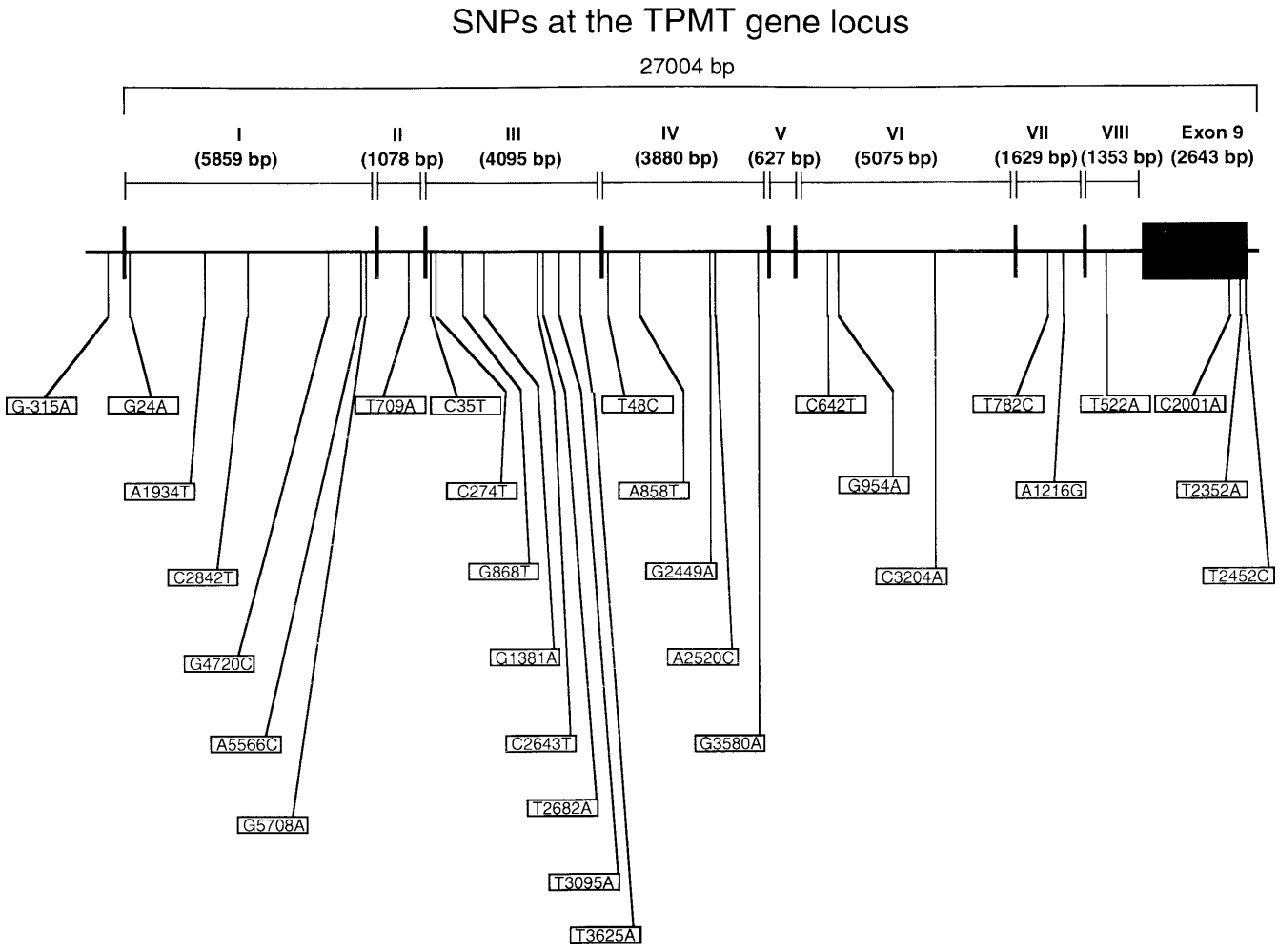

Fig. 1. Genomic structure of the *TPMT* gene. *Black bars* depict exons, and *Roman numerals I–VIII* denote introns. *Arabic numbers* indicate the length of each intron, in bp. *SNP*, Single-nucleotide polymorphism

comparison to their frequencies in white and black subjects (Collie-Dugnid et al. 1999). Furthermore, differences (up to fivefold) in levels of TPMT activity occur among Americans who carry two copies of the wild-type *TPMT* allele, indicating that other genetic factors might be involved.

In the present study, we determined the entire genomic structure of the *TPMT* gene and undertook screening for SNPs outside coding regions. The genomic structure reported here is different from that reported elsewhere (Szumlanski et al. 1996) (e.g., nine exons instead of ten, and 27 kb instead of 32 kb), but, because we PCR-amplified and sequenced all intronic regions using genomic DNAs from multiple individuals, we have confidence in our results concerning the number of exons and the size of each intron.

The frequency of SNPs (1 in every 907 bp) detected in our study was lower than we anticipated on the basis of earlier studies, probably because we examined fewer chromosomes. However, the 30 markers reported here should be useful for seeking associations between alleles of the *TPMT* gene and side-effects caused by drugs whose metabolic fates are affected by this enzyme. Our results underscore the potential usefulness of SNP genotyping in clinical settings to determine safe dosages of aromatic and heterocyclic sulfhydryl compounds. We hope our data can contribute to further investigations of the *TPMT* gene and to clinical research.

References

Ameyaw MM, Collie-Duguid ES, Powrie RH, Ofori-Adjei D, McLeod HL (1999) Thiopurine methyltransferase alleles in British and Ghanaian populations. Hum Mol Genet 8:367–370

Collie-Duguid ES, Pritchard SC, Powrie RH, Sludden J, Collier DA, Li T, McLeod HL (1999) The frequency and distribution of thiopurine methyltransferase alleles in Caucasian and Asian populations. Pharmacogenetics 9:37–42

Elion GB (1989) The purine path to chemotherapy. *Science* 244:41–47

Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE (1999) Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. Hum Mol Genet 8:371–376

Krynetski EY, Evans WE (1998) Pharmacogenetics of cancer therapy: getting personal. Am J Hum Genet 63:11–16

Krynetski EY, Krynetskaia NF, Yanishevski Y, Evans WE (1995a) Methylation of mercaptopurine, thioguanine, and their nucleotide metabolites by heterologously expressed human thiopurine S-methyltransferase. Mol Pharmacol 47:1141–1147

Krynetski EY, Schuetz JD, Galpin AJ, Pui C-H, Relling MV, Evans WE (1995b) A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. Proc Natl Acad Sci USA 92:949–953

Krynetski EY, Tai H-L, Yates CR, Fessing MY, Loennechen T, Schuetz JD, Relling MV, Evans WE (1996) Genetic polymorphism of thiopurine S-methyltransferase: clinical importance and molecular mechanisms. Pharmacogenetics 6:279–290

McLeod HL, Lin J-S, Scott EP, Pui C-H, Evans WE (1994) Thiopurine methyltransferase activity in American white subjects and black subjects. Clin Pharmacol Ther 55:15–20

McLeod HL, Krynetski EY, Willimas JA, Evans WE (1995) Higher activity of polymorphic thiopurine S-methyltransferase in erythrocytes from neonates compared to adults. Pharmacogenetics 5:281–286

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

Otterness D, Szumlanski C, Lennard L, Klemetsdal B, Aarbakke J, Park-Hah JO, Iven H, Schmiegelow K, Branum E, O'Brien J, Weinshilboum R (1997) Human thiopurine methyltransferase pharmaco genetics: gene sequence polymorphisms. Clin Pharmacol Ther 62:60–73

Otterness DM, Szumlanski CL, Wood TC, Weinshilboum RM (1998) Human thiopurine methyltransferase pharmacogenetics. Kindred with a terminal exon splice junction mutation that results in loss of activity. J Clin Invest 101:1036–1044

Szumlanski C, Otterness D, Her C, Lee D, Brandriff B, Kelsell D, Spurr N, Lennard L, Wieben E, Weinshilboum R (1996) Thiopurine methyltransferase pharmacogenetics: human gene cloning and characterization of a common polymorphism. DNA Cell Biol 15: 17–30

Tai HL, Krynetski EY, Yates CR, Loennechen T, Fessing MY, Krynetskaia NF, Evans WE (1996) Thiopurine S-methyltransferase deficiency: two nucleotide transitions define the most prevalent mutant allele associated with loss of catalytic activity in Caucasians. Am J Hum Genet 58:694–702

Weinshilboum RM, Sladek SL (1980) Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet 32:651–662