METHODS

RT-PCR Permits Simultaneous Genotyping of Thiopurine S-methyltransferase Allelic Variants by Multiplex Induced Heteroduplex Analysis

Nigel Wood,1* Adrian Fraser,2 Jeffrey Bidwell,2 and Graham Standen3

1Department of Pathology and Microbiology, School of Medical Sciences, University of Bristol, Bristol, UK; 2Department of Pathology and Microbiology, University of Bristol, Homoeopathic Hospital Site, Cotham, Bristol, UK; 3Molecular Haematology Unit, Department of Haematology, Bristol Royal Infirmary, Bristol, UK

Communicated by Linda Tyfield

Thiopurine-based drugs are a widely prescribed group of medications. Their tolerance and effectiveness is dependent on an individual’s ability to metabolize these compounds. An essential enzyme for the metabolism of these drugs is thiopurine S-methyltransferase (TPMT), whose activity is subject to genetic variation. Genotyping of the most frequent allelic variants in TPMT affords an extremely accurate prediction of the three clinical phenotypes: high, intermediate, and low enzyme activity. One constraint of most genotyping methods is the inability to demonstrate physical linkage between two sequence variants that occur in different exons, e.g., c.460G>A and c.719A>G, which give rise to TPMT*3, the most common defective allele in Caucasian populations. Using mRNA/cDNA as a template enables analysis of both sequence variants in a single assay. This approach could be applicable to other genes where allelic variation (in-cis and in-trans) is due to alterations in different exons. Induced heteroduplex generator analysis has previously been shown to discriminate in-cis and has also been suitable for multiplexing. In this method we have exploited both these features and for the first time have applied them to a RT-PCR analysis. The primary reagent developed allows unequivocal resolution of TPMT*3A and the alleles carrying the c.719A>G allelic variant, TPMT*3C, as well as the silent alteration c.474T>C. The TPMT*3B variant has not been observed. A secondary reagent, which can be multiplexed, identifies the TPMT*2 allele. Hum Mutat 24:93–99, 2004. © 2004 Wiley-Liss, Inc.

KEY WORDS: TPMT; thiopurine methyltransferase; RT-PCR; IHG, induced heteroduplex generator; genotyping; sequence specific multiplex PCR

DATABASES:
TPMT– OMIM:187680; GDB: 209025; GenBank: NM_000367.1 GeneCards: GC06M018186

INTRODUCTION

The enzyme thiopurine methyltransferase (TPMT; MIM#187680) catalyzes the metabolism of the widely administered thiopurine-based group of drugs that includes 6-mercaptopurine, azathioprine, and 6-thioguanine [Weinshilboum and Sladek, 1980]. These agents are used in the treatment of acute leukemia and as immunosuppressants for the management of organ transplantation and autoimmune disorders (e.g., inflammatory bowel disease, atopic dermatitis, and rheumatoid arthritis) [Wells et al., 1994; Kerstens et al., 1995; Tan et al., 1997; Meggitt and Reynolds, 2001; Schwab et al., 2001; Thervet et al., 2001]. Clinical toxicity of these drugs particularly involves the bone marrow and gut [Singh et al., 1989] and has been attributed to variation in cytosolic activity of the enzyme. This variation of functional enzyme results from a number of sequence variants that have been identified in the reading frame of the TPMT gene.

Three common sequence variants, which account for a high proportion of allelic variation in TPMT activity, have been reported (all sequence variants described are numbered from the A residue of the ATG translation initiation start site). The first to be identified was a c.238G>C transversion in exon 5 [Krynetski et al., 1995], which results in substitution of proline for alanine at codon 80 (p.A80P). The expressed activity of this allele (designated TPMT*2) is approximately 5% of the wild type allele, TPMT*1 [Tai et al., 1996].

Received 27 August 2003; accepted revised manuscript 5 March 2004.

*Correspondence to: Dr. N.A.P.Wood, Department of Pathology & Microbiology, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, UK. E-mail: nigel.wood@bristol.ac.uk

The induced heteroduplex generator technology is covered by British Patent No. 2280266.

DOI 10.1002/humu.20057
Published online in Wiley InterScience (www.interscience.wiley.com).
The most common variant allele is TPMT*3A, which consists of a c.460G → A transition, p.A154T in exon 7 linked in cis with an c.719A → G transition, p.Y240C in exon 10 [Szumlanski et al., 1996; Tai et al., 1996]. This allele has a frequency of 3.9 to 5.7% in European Caucasians, accounts for 72 to 78% of mutant TPMT alleles in this population [Otterness et al., 1997; Rossi et al., 2001], and has no measurable enzyme activity [Tai et al., 1996]. Less commonly, the c.460G → A (TPMT*3B) and c.719A → G (TPMT*3C) sequence variants are found in isolation.

The precise biological role of TPMT enzyme is unknown but its clinical relevance is due to the enzyme’s capacity to methylate 6-mercaptopurine (6-MP) and the catalyst thionine monophosphate (tIMP) [Coulthard et al., 2000]. In the presence of high enzyme activity, these catalytic pathways act as competitive regulators in the formation of 6-thioguanine nucleotides (TGNs) [Lennard et al., 1997]. In individuals with low TPMT activity, this methylation fails to occur and TGNs accumulate [Lennard et al., 1997]. Excessive incorporation of TGNs during DNA synthesis [Lennard, 2001] is cytotoxic and accounts for the severe bone marrow suppression seen in TPMT deficiency [Lennard et al., 1989]. The observed trimodal distribution of phenotypes correlates to high activity (homozygous for the wild type allele, TPMT*1), intermediate activity (heterozygous for a mutant allele), and low activity (homozygous for a mutant allele).

It is now generally recommended that patients should be screened for TPMT deficiency prior to administration of thiopurine agents. Reduced dose requirements can be predicted in patients with intermediate or low TPMT activity and the risk of severe bone marrow aplasia can be eliminated [Lennard et al., 1997; Standen and Wood, 2001]. A phenotypic analysis that measures total enzyme activity directly in red cells is the most accurate and clinically informative result. The current routine method involves incubation with an isotopic drug analog followed by extraction with organic solvents [Coulthard et al., 2000]. A more recent adaptation is a HPLC assay, which does not require radioactive isotopes [Menor et al., 2001]. However, this direct approach can be subject to biological anomalies. For example, children presenting with acute lymphoblastic leukemia (ALL) have a predominantly older population of red cells which have been shown to have lower TPMT activity [Lennard et al., 2001]. Conversely, falsely elevated activity may be found if blood transfusions [Schwab et al., 2001] are given prior to sampling. Currently, the phenotypic measurement of TPMT activity is carried out routinely in only a small number of specialized laboratories in the UK.

Genotyping involving DNA analysis of the most frequent TPMT sequence variants is able to predict enzyme activity and is widely available. A genotype-directed dosing strategy is also cost effective [Marra et al., 2002]. However, there are a number of drawbacks to DNA analysis methods that are currently used. Restriction endonuclease digestion of PCR products is the most frequently used method to identify the various specific sequence variants [Yates et al., 1997]. However, partial or incomplete digestion resulting from incompatibilities between restriction enzyme and PCR buffers has recently been shown to cause frequent aberrant results and misdiagnosis of the genotype [Brouwer et al., 2001]. To identify the *3B allele, Schutz et al. [2000] and Brouwer et al. [2001] have used cloning of cDNA followed by sequencing or restriction enzyme digestion.

We describe here a novel adaptation of the DNA induced (or universal) heteroduplex generator (IHG) technology that has previously been applied to sequence variant analysis of PAH, β-globin, and CFTR genes [Wood and Bidwell, 1996]. As a conformational method, it has the capacity to identify multiple sequence variations representing a single alleles, which is a unique advantage over most sequence variation detection systems [Bolla et al., 1999; Turner et al., 2000; Wood et al., 2001; Li et al., 2002]. Earlier applications of this technology has been applied only to genomic DNA templates but the use of mRNA templates for reverse transcription PCR (RT-PCR) has permitted the demonstration of physical linkage (haplotype) between the two sequence variants present in the TPMT*3A allele (Fig. 1).

A further refinement has been to multiplex the heteroduplex analysis [Barbax et al., 2000] by the inclusion of a second reaction to detect the TPMT*2 allele. Potentially, this system enables detection of the most frequent and clinically relevant sequence variants within the TPMT gene in a single assay. The possible effectiveness of such a single genotyping assay was demonstrated by a recent study of a European population [Rossi et al., 2001]. In this study, both the TPMT genotypes and TPMT phenotypes were assayed, only TPMT*2 and TPMT*3 allelic variants were observed, with a concordance of 97% between the two assay methods.

**MATERIALS AND METHODS**

**Primer Design and IHG Construction**

The presence of an intronless TPMT pseudogene [Lee et al., 1995] necessitated the design of 3’ sequence-specific primers to the TPMT cDNA sequence [Lee et al., 1995] (Table 1). This eliminated the possibility of amplifying contaminating, residual, genomic DNA in the mRNA preparations. Alignments of the GenBank submissions U11424.1 (pseudogene) [Lee et al., 1995], AF021877.1 (partial mRNA sequence) [Yates et al., 1997], NM_000367.1 (mRNA sequence) [Otterness et al., 1997], and AF019369.1 (promoter exon 10 and intron) [Krynetski et al., 1997] were constructed with the Multiple Alignment Construction and Analysis Workbench (MACAW; Version 2.0) [Schuler et al., 1991]. Nucleotide numbering for DNA variants uses +1 as the A of the initiation codon.

The TPMT*3 IHGs were constructed from four head-to-tail, overlapping oligonucleotide longmers [Wood and Bidwell, 1996; Bolla et al., 1999]. Longmer 1 contained the identifier for the c.460G → A transition, and longmer 4 contained the identifier for c.719A → G transition. A series of empirical refinements to the longmers containing the identifiers determined the most sensitive and definitive IHG sequence, designated TPMT*3 IHG(3) (details available upon request). The coding sequence contained an insert of three guanine residues immediately upstream of the
c.460G>A polymorphism and an insert of three cytosine residues immediately upstream of the c.719A>G polymorphism.

The design of TPMT*2 IHG was dictated by the parameters established for the TPMT*3 IHG(3) analysis. cDNA sequence-specific primers were designed that had similar Tms to that of TPMT*3 primers (Table 1). The IHG, designated TPMT*2 IHG(1), was amplified from a single longmer with an insert of four adenosine residues immediately downstream from the c.238G>C transversion. The resulting IHG, having an apparent increase in molecular size 200 to 300%, would migrate just ahead of the homoduplex of the TPMT*3 products on a 15% gel (see Fig. 2).

Reverse Transcription and PCR Parameters

Pilot assays, to demonstrate primer specificity, were done with a two-step procedure, cDNAs for PCR being generated from total RNA preparations using MMLV reverse transcriptase (Promega UK, Southampton, UK) and an oligo dT primer. For convenience and cost effectiveness, subsequent RT-PCRs were routinely performed using a one-step commercially available kit, the Qiagen® OneStep® kit (www.qiagen.com). This contains a mixture of two reverse transcriptases, Omniscript and Sensiscript Reverse Transcriptases, which are recombinant heterodimeric enzymes expressed in E. coli, and a modified Taq polymerase (HotStarTaq DNA Polymerase) that requires heat activation.

The manufacturer’s protocol was adapted for multiplex RT-PCRs in 20 μL reactions (Table 2). RT-PCR reactions were performed on an MJ PTC100 thermal cycler (G.R.I. Braintree, Essex, UK) or Robocycler Gradient 96 (Stratagene, www.stratagen.com), using the following touchdown PCR [Hecker and Roux, 1996] program: 30 min at 50°C (reverse transcription); 15 min at 95°C (inactivation of reverse transcriptases and activation of DNA polymerase and DNA denaturation); 12 cycles of 95°C for 20 sec, 56°C for 15 sec, and 72°C for 20 sec; 16 cycles of 95°C for 30 sec, 53°C for 25 sec, and 72°C for 30 sec; 7 cycles of 95°C for 45 sec, 51°C for 45 sec, and 72°C for 45 sec (5 sec was added to each step of the program when performed on the Robocycler); and a final extension of 72°C for 5 min. Then 6.5-μL aliquots of each individual reaction were qualitatively assessed on a 2% agarose gel, 0.5x TBE running buffer (containing 0.5 μg/mL ethidium bromide).

Samples and Verification of Genotype

Total RNA was extracted from whole blood samples from patients with a wide range of hematological disorders. Extraction was via the single step method TRIzol® (Invitrogen Ltd., Paisley, UK) [Chomczynski and Sacchi, 1987] or PUREscript™ isolation kits (Flowgen Instruments Ltd, Lichfield, UK), both used according to manufacturers' instructions. The third variant pattern was resolved by a restriction digest with the enzyme Ase I (New England Biolabs, Hitchen, UK).

Dual IHG Construction and Multiplex Analysis

The efficacy of the individual IHGs having been verified, a combination IHG was constructed to ensure that both IHGs were at an equimolar concentration for multiplex amplification. TPMT*2 reverse (exon 5) and TPMT*3 forward (exons 6 and 7)
primers were synthesized with complementary (ACGT)\textsubscript{16} tails. Both TPMT\textsuperscript{*2} IHG(1) and TPMT\textsuperscript{*3} IHG(3) were amplified with the requisite modified primers; stocks of purified individual modified IHG amplicons were coamplified with primers (TPMT\textsuperscript{*2} forward, exon 4; and TPMT\textsuperscript{*3} reverse, exon 10). This reaction, was catalyzed by a proofreading DNA polymerase, BIO-X-ACT-Short (Bioline Ltd., London, UK). The linked IHG amplicon was gel purified [Bolla et al., 1999] and cloned (Qiagen PCR Cloning plus Crawley, Sussex, UK) Plasmid DNA, for generation of IHG stocks by amplification, was isolated using the method of Holmes and Quigley [Sambrook et al., 1989].

PCR of the combined IHGs were performed either on a MJ PTC100 thermal cycler, Robocycler Gradient 96, or an Idaho Technology Rapid Cycler (Biogene, Kimbolton, UK). The linked IHG amplicon was gel purified [Bolla et al., 1999] and cloned (Qiagen PCR Cloningplus Crawley, Sussex, UK) Plasmid DNA, for generation of IHG stocks by amplification, was isolated using the method of Holmes and Quigley [Sambrook et al., 1989].

PCR of the combined IHGs were performed either on a MJ PTC100 thermal cycler, Robocycler Gradient 96, or an Idaho Technology Rapid Cycler (Biogene, Kimbolton, UK). The linked IHG amplicon was gel purified [Bolla et al., 1999] and cloned (Qiagen PCR Cloningplus Crawley, Sussex, UK) Plasmid DNA, for generation of IHG stocks by amplification, was isolated using the method of Holmes and Quigley [Sambrook et al., 1989].

Heteroduplex Generation and Analysis

Induced heteroduplex generation (otherwise referred to as cross-matching) was achieved by denaturing 12 \textmu L of multiplexed cDNA amplicons with 10 \textmu L of multiplexed IHG amplicons, followed by slow controlled cooling. Induced heteroduplex analysis requires the reannealed DNA molecules to be resolved on a molecular sieving gel matrix. The most informative resolutions were achieved on polyacrylamide gels with a low degree of cross-linkage, 2.6\% (acrylamide/methylene bisacrylamide solution; 37.5:1 ratio).

Heteroduplex electrophoresis was carried out on a C.B.S. triple-wide minigel system, which can carry two 8 x 30 cm, 34- or 50-lane gels (Scientific Laboratory Supplies, Wilford, UK), potentially providing sufficient lanes for the analysis of one 96-well PCR plate.

Satisfactory resolution of both TPMT\textsuperscript{*3} IHG3 and TPMT\textsuperscript{*2} heteroduplex bands, from the multiplex reactions, was obtained on 15\% Protogel\textsuperscript{TM} (National Diagnostics Ltd., Hull, UK) 2.6\% cross-linkage gels containing 1 x TBE, 1 x TBE running buffer, circulating coolant at 10°C with the following electrophoresis conditions: constant voltage (300 V) for 135 min. Gels were stained post-electrophoresis with SYBR Green (7.5 \textmu L of 10,000 stock per 100 mL 1 x TBE buffer; Molecular Probes, Leiden, The Netherlands). Images for all analyses were acquired using a Kodak EDAS 240 system (Anachem, Luton, UK).
RESULTS

Amplification of eight individual genomic DNA control samples was used to confirm the specificity of primer designs. Both sets of TPMT cDNA-specific primers were demonstrated to be refractory to the amplification of genomic DNA (data not shown).

Since pilot experiments using a two-step RT-PCR system (data not shown) showed TPMT mRNA to be constitutively expressed with little intersample variation, subsequent analyses were performed using the one-step system.

TPMT IHG analysis was performed on 47 clinical samples from individuals presenting at the Bristol Royal Infirmary with various hematological disorders. A series of gel optimizing experiments determined that the most informative separation of the heteroduplex bands was obtained on a 15% Protogel. Using the described parameters and this percentage PAGE, the TPMT*3 IHG(3) gave clearly-defined differentiation between the previously identified TPMT*3A/*3C alleles. Also, a third heterozygous pattern (Fig. 1; lanes 2, 12, 13, and 16) was observed; a simple RFLP analysis with the restriction enzyme Ase I showed that this pattern was caused by the presence of the silent sequence variant c.474T>C. Multiplexing of both TPMT*3 and TPMT*2 PCR reactions required considerable optimization: this was carried out along the guidelines established by Henegariu et al. [1997]. Balanced amplification of both IHGs was obtained by increasing the number of units of Taq polymerase added to each reaction.

With the multiplex gel analysis, no variant patterns were observed with the TPMT*2 IHG(1) and the 47 clinical samples; to verify the efficacy of the reagent, an amplification from a full length TPMT*2 cDNA clone was performed. When cross-matched with TPMT*2 IHG, this amplification gave a discrete banding pattern that was clearly distinguishable from that given by TPMT*1 (Fig. 2, lane 18). In the multiplex assay, it gave TPMT*1 type pattern with TPMT*3 IHG(3).

DISCUSSION

Approximately 1 out of 300 individuals have TPMT deficiency and are at high risk of developing severe hemopoietic toxicity following administration of thiopurine drugs.

Although screening for the inherited defect is recommended prior to treatment, TPMT enzyme assays are relatively complex, labor intensive, and are not widely available in routine chemical pathology laboratories. This has encouraged the development of a variety of DNA-based methods, which have been designed to detect common sequence variants in the TPMT gene associated with impaired enzyme function. In this report, we describe an advancement of the proven IHG analytical method whereby an amplicon from a cDNA is analyzed. This permits the simultaneous determination of linked sequence variants about 3.0 kb apart, thereby establishing a haplotype in a single reaction. This method enables rapid characterization of the common mutant TPMT alleles *2, *3A, *3B, and *3C.

In a recent study of 18 subjects with very low or absent erythrocyte TPMT activity, these alleles accounted for all but one of the 36 mutant haplotypes identified [Otterness et al., 1997]. Our results are concordant with those seen in other regional studies within the UK where the population has multiple ethnicities but is predominately Caucasian. Comparison of UK Caucasians with African and Asian population groups [Ameyaw et al., 1999; Collie-Duguid et al., 1999; McLeod et al., 1999] show that all common sequence variants are found in the Caucasians, but only TPMT*3C was present in Chinese and African populations, though a TPMT*3A allele was observed in a UK Southwest Asian cohort. One of the confirmed TPMT*3C genotyped individuals was of African-Caribbean origin; given the predominance of this allele in this major human group, this result was not unexpected. The frequency of the silent polymorphism has been shown to be as high as 21% in some clinical cohorts [Otterness et al., 1997], and the Bristol cohort was shown to have a similar frequency.

Homozygotes of the polymorphic variants and the TPMT*3B allele were not seen in this small cohort. Nevertheless, based on previous studies using IHG technology, and having established unique banding patterns for two of the variant alleles, homozygotes could be readily analyzed in future studies. It would also be predicted that the TPMT*3B allele could be identified by a unique banding pattern. As we have already demonstrated with the detection of one of the alleles (TPMT*3C) that occurs due to a single alteration, it would be expected that the IHG would form a further unique conformation in the presence of the other allele that resulted from an alternative single alteration (TPMT*3B).

A number of TPMT genotyping assays have been developed for the new sequence variant analysis platforms now available; these require sophisticated pieces of apparatus such as automated DNA sequencing systems [Alves et al., 2000], DHPLC [Hall et al., 2001], or fluorescence emission based real-time PCR [Schutz et al., 2000]. Such apparatus has high initial capital expenditure and high consumable and maintenance costs. Even with these systems, there are potential flaws that could compromise the clinical efficacy of the analysis. For example, methods that rely on DNA duplex melting must be able to distinguish between the silent alteration c.474T>C in exon 7, which is only 14 nucleotides from the c.460G>A site [Spire-Vayron de la Moureyre et al., 1998].

As we described in previous applications of IHG technology [Wood and Bidwell, 1996], there are only minimal capital expenditures associated with the method; the gel-based format permits flexibility with regard to sample numbers processed (10–100 samples) and the necessary apparatus is often already available in the laboratory. Due to the generation of unique conformational structures, the presence of the silent polymorphism cannot cause misinterpretation of the analysis. Furthermore, the
method is not subject to inaccurate analysis due to partial or inhibited RFLP digests [Brouwer et al., 2001], which are the most frequently used of the low-tech TPMT genotyping methods. Multiplexing of IHG analysis has previously been described [Bowen et al., 1998; Barboux et al., 2000], as has haplotype determination [Bolla et al., 1999; Wood et al., 2001; Li et al., 2002]. This application exploits both of these facets, but this is the first report of IHG analysis from a mRNA/cDNA template. The method described here for TPMT genotyping analysis could be readily assimilated by laboratories that have molecular/PCR experience, but that are unable to offer a cellular, phenotypic analysis.

Apart from long-range PCR (LR-PCR) with intramolecular ligation [McDonald et al., 2002], all previously described TPMT genotyping methods fail to address the problem of identifying multiple sequence variants on a single chromosome, such as the cis linkage between the polymorphic sites c.460G>A and c.719A>G, which comprise the TPMT*3A allele. Using genomic DNA templates, the distance between these two sites is greater than 3.0 kb, which is outside the range of discrimination for the majority of current systems and would require excellent DNA preparations. This is clinically important because such methods fail to distinguish between compound heterozygous TPMT*3B/*3C and heterozygous TPMT*3A/*1 haplotypes. Linkage in cis has normally been demonstrated only by cloning and subsequent sequencing (and now by the recently described LR-PCR method). However, by using mRNA as the starting material, the IHG technique provides a genotyping system that avoids the potential problem of pseudoheterozygosity [Schutz et al., 2000].

Routine use of mRNA in molecular diagnostics laboratories has become possible because of recent commercial developments. The RT-PCR kit described proved to be extremely robust; reactions were no more complex to set up than routine PCRs, and combining the two steps permitted a full analysis within 8 hr. With the flexibility of simple gel-based multiplex IHG analysis, and the commercial innovations for collection and extraction of mRNA, this method of genotyping could provide the level of sensitivity required for a feasible alternative to the TPMT enzyme activity assay for clinical screening. With the implementation of this method of genotyping, only specific cases, in which a definitive measure of enzyme activity is required, would necessitate phenotypic analysis.

ACKNOWLEDGMENTS

We thank Dr. Eugene Krynetski for supplying the TPMT*2 full-length cDNA clone and for his appraisal of this manuscript, and we thank Doris Culpan for nucleotide sequencing.

REFERENCES


Bowen DJ, Bowley S, John M, Collins PW. 1998. Factor V Leiden (G1691A), the prothrombin 3’-untranslated region variant (G20210A) and thermolabile methylenetetrahydrofolate reductase (C677T): a single genetic test genotypes all three loci—determination of frequencies in the S. Wales population of the UK. Thromb Haemost 79:949–954.


