UNIVERSITY STUDY OF MILAN

Graduate School in Pharmacological Sciences

Department of Medical Pharmacology
PhD in Medical Pharmacology, Chemotherapy and Toxicology
XXIV Cycle

PhD Thesis

Towards predictive pharmacogenetic profiling for azathioprine treatment: characterization of SNPs in relevant drug metabolism genes

BIO/14 - Pharmacology

Dr. Lara Colleoni
R08308

TUTORS:
Dr. Pia Bernasconi

Prof. Alberto Panerai

CO-ORDINATOR:
Prof. Alberto Panerai

Academic year 2010 – 2011

Ai miei genitori
e
a Daniele

INDEX

SUMMARY

Azathioprine (AZA) is one of the most widely used immunosuppressant drugs for the treatment of many inflammatory diseases, in particular autoimmune diseases where it is often administered as a steroid-sparing drug. It blocks T- and B-cell proliferation through incorporation of its active metabolites, 6-thioguanine nucleotides, into DNA. AZA is mainly catabolised through the thiopurine S-methyltransferase (TPMT) and xanthine oxidase (XO) pathways. While 70% of patients respond to AZA, 30% do not respond or show intolerance to the drug. Variants of genes coding for TPMT or XO, or other gene members of either pathway, which result or not in mutated enzymes, are believed to influence individual responses to AZA. With the aim of uncovering possible genotype-phenotype correlations that would help personalize AZA treatment through response-specific SNP profiles, we have analyzed *TPMT* and *XDH* (the gene coding for XO) for SNPs in 71 Italian patients in the context of their response to AZA (intolerant patients, n = 25; unresponsive patients, n = 16; and responsive patients, n = 30). We confirm the presence of known intronic and exonic *TPMT* and *XDH* polymorphisms not correlating with particular AZA responses. We identified two new intronic polymorphisms, G420-4A in *TPMT* and T652-21A in *XDH*, each in one intolerant patient. Whether or not intolerance is related to these mutations in the patients must be further elucidated. In *XDH*, we detected two novel non-synonymous mutations (c.G1004A, p.Arg335His; c.C2891T, p.Thr964Ile), one new synonymous mutation (c.G1194A, p.Leu398Leu), and one novel non-synonymous polymorphism (c.C1167T, p.Ala556Val). The missense mutation, p.Arg335His, is likely to have an effect on the structural conformation of the FAD-binding domain, thereby destabilizing the protein structure. This could therefore lead to a decreased XO activity, for example through increased protein degradation or through deficient binding of FAD, hence reduced efficiency of the enzyme function, and thereby intolerance to AZA as observed. The other non-synonymous mutation, p.Thr964Ile, occurred in the molybdopterin domain; it was found in one responsive patient, suggesting that this mutation did not modify the structure of the domain sufficiently to affect the enzymatic activity. The non-synonymous polymorphism, p.Ala556Val, causes a change from Ala to Val at position 556 in the connection segment between the FAD-binding and molybdopterin domains, which is unlikely to have a significant effect on the protein. This *XDH* polymorphism was found in one intolerant, one unresponsive, two responsive patients, and 3/100 healthy controls; it is therefore unlikely to be of significance in

pharmacogenetic profiles predictive of AZA responses. The silent mutation, p.Leu398Leu, identified in an intolerant patient should not cause functional impairment of XO.

Linkage disequilibrium (LD) of *TPMT* SNPs and haplotype analysis thereof demonstrated a new haplotype designated *TPMT**3E; it comprises the previously reported mutations of the *TPMT**3A allele associated with intolerance to AZA and the intronic T140+114A SNP. *TPMT**3E was detected in four of the 25 AZA-intolerant patients and was not observed in unresponsive or responsive patients. The association of *TPMT*3E* with AZA intolerance, and its frequency, must be ascertained in larger, ethnically different cohorts. Nevertheless, in view of the highly significant association (Psim = 0.037) between *TPMT**3E and AZA intolerance in our study, this haplotype should be taken into account when considering AZA treatment.

LD analysis of *XDH* identified four different haplotype blocks, one of which was significantly associated with intolerance in our cohort (Psim = 0.017). This block includes five SNPs, one intronic and four located in the 3' untranslated regions (3' UTRs); these were previously described as single SNPs, but were not analyzed in the context of response to AZA. It is unclear how this haplotype results in intolerance to AZA. One possibility is that the SNPs affect the regulation of protein expression through alteration of the target sites for microRNAs that interact with the 3' UTRs to regulate the expression of mRNAs. This could therefore lead to AZA intolerance in our patients, through in an increase in down-regulation of *XDH* mRNA and thereby XO expression. In this work, we have demonstrated new haplotypes that should be taken into consideration in pharmacogenetic profiling for AZA. In particular, SNPs in the *XDH* have been poorly investigated thus far in the context of response to AZA; the new *XDH* haplotype is of major interest in the establishment of pharmacogenetic profiles that will permit prediction of the type of response to AZA, in particular to prevent life-threatening side effects. It should be further studied in the context of its association with other response-defining haplotypes or SNPs of *TPMT* and other AZA metabolism pathway genes.

INTRODUCTION

1. AZATHIOPRINE

1.1. History

Azathioprine (AZA) was synthesised in 1957 by Elion and Hitchings.[1] Azathioprine is among the oldest pharmacologic immunosuppressive agents in use today. Initially developed as a long-lived prodrug of 6-mercaptopurine (6-MP), it was quickly found to have a more favourable therapeutic index.[2] Initially evaluated for the treatment of leukaemia in the 1950s, AZA was introduced in the 1960s as immunosuppressant for organ transplantation. The use of thiopurine has also been reported for inflammatory bowel disease, such as Crohn's disease and ulcerative colitis,[3] haematological malignancies, rheumatologic diseases,[2] and autoimmune diseases.[4-6]

AZA (**Figure 1a**) and 6-MP (**Figure 1b**) have been used in the clinic for more than five decades. Inhibition of purine nucleotide biosynthesis and eventually suppression of DNA and RNA synthesis, and down-regulation of B and T cell proliferation have been suggested as major therapeutic mechanisms.[7] The immunosuppressive action of AZA is delayed, starting 3-6 months after treatment is initiated.[4]

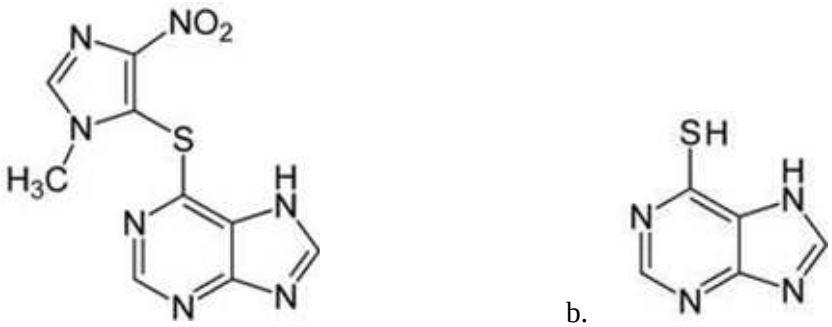


Figure 1. Chemical structure of (a) azathioprine and (b) 6-mercaptopurine

1.2. Mechanism of action

Azathioprine is a prodrug, a pharmacologically inactive chemical derivative that can be used to transiently alter the physicochemical properties of drugs, to increase their usefulness and/or to decrease associated toxicity. Generally, prodrugs can be converted in

vivo, enzymatically or non-enzymatically, into active drug molecules to exert a therapeutic effect.[8]

After oral administration and absorption, approximately 90% of the prodrug AZA undergoes conversion to 6-MP, the pharmacological molecule, by non-enzymatic binding of sulphydryl-containing compounds such as glutathione (GSH) or cysteine that are present in every mammalian cell.[7] As shown by other studies, another mode of biotransformation of AZA involves enzymatic conjugation with GSH to form 6-MP, catalyzed by glutathione S-transferase (GST). GST consumes GSH, which is normally present in abundance in hepatocytes[9] (**Figure 2**).

AZA is believed to act through the insertion into DNA and RNA of its metabolic products, thioguanine nucleotides (TGNs). These are produced through the metabolic pathway which converts 6-MP to 6-thioinosine 5'-monophosphate (TIMP), TIMP to thioxanthine monophosphate (TXMP), and TXMP to thioguanosine monophosphate (TGMP), through the action of hypoxanthine-guanine phosphoribosyl transferase (HGPRT), inosine monophosphate synthetase (IMPDH), and guanosine monophosphate synthetase (GMPS), respectively (**Figure 2)**. TGMP is then metabolised to form di- and tri-phosphate TGNs. Thiopurine S-methyltransferase (TPMT), aldehyde oxidase (AO), and xanthine oxidase (XO) compete with HGPRT, which initiates the metabolism of 6-MP to TGNs, for the initial metabolism of 6-MP. TPMT can act not only on 6-MP, but also on two intermediates of the HGPRT pathway, TIMP and TGMP. As substrates for TPMT, 6-MP and TGMP form methylmercaptopurine (MeMP) and methyl thioguanosine monophosphate (MeTGMP), respectively,[10] whereas TIMP is converted to S-methyl-thioinosine 5'-monophosphate (MeTIMP), a strong inhibitor of *de novo* purine synthesis believed to contribute significantly to the cytotoxic action of 6-MP.

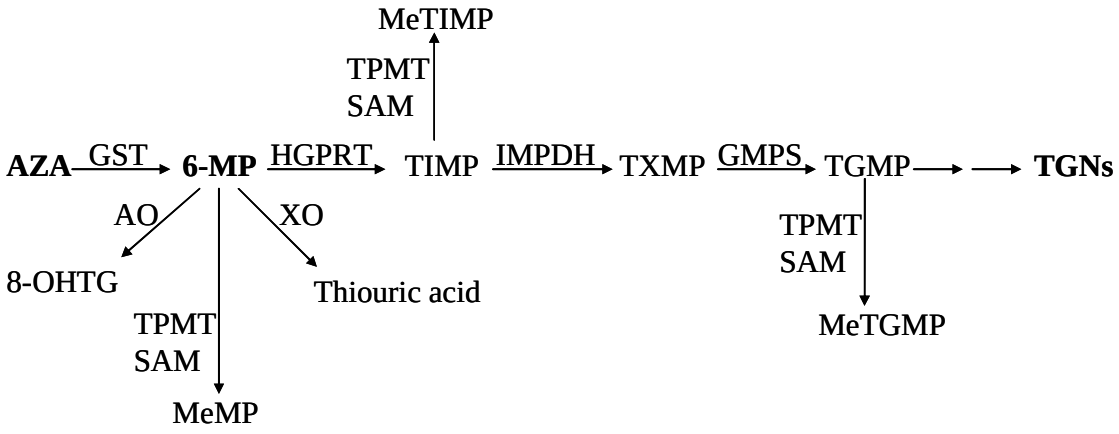


Figure 2. Metabolism of AZA. The various steps leading to the formation of thioguanine nucleotides (TGNs) are catalyzed by glutathione S-transferase (GST), hypoxanthine guanine phosphoribosyltransferase (HGPRT), inosine monophosphate dehydrogenase (IMPDH), guanosine monophosphate synthase (GMPs). The oxidation is catalyzed by xanthine oxidase (XO) and aldehyde oxidase (AO); the methylation is carried out by thiopurine methyltransferase (TPMT) using the methyl donor S-adenosyl-L-methionine (SAM). The compounds produced in the various metabolic steps are: 8-hydroxythioguanine (8-OHTG); methyl-mercaptopurine (MeMP); thioinosine-monophosphate (TIMP); methyl-thioinosine-monophosphate (MeTIMP); thioxanthine-monophosphate (TXMP); thioguanine-monophosphate (TGMP) and methyl- thioguanine-monophosphate (MeTGMP).

The TGN metabolites of 6-MP act as purine antagonists and inhibit DNA, RNA, and thereby protein synthesis, inducing cytotoxicity and immunosuppression. More recently, AZA was also shown to interfere at the protein level itself. Tiede et al. showed that *in vitro* stimulation of primary human T lymphocytes in the presence of AZA or 6-MP results in an increased percentage of apoptotic cells.[7] 6-MP was shown to induce apoptosis by converting the CD28-mediated co-stimulatory signalling implicated in T-cell activation into an apoptotic signal, through direct interaction of its 6-thioguanine triphosphate metabolite with Rac1, a small GTPase whose activation upon CD28 costimulation appears to control pathways known to prevent T-cell apoptosis. Binding of 6-thioguanine triphosphate to Rac1 apparently suppresses the activation of Rac1 target genes leading to a mitochondrial pathway of apoptosis.[7]

1.3. Side effects

Adverse effects may occur in 15 to 30% of patients treated with AZA. These side effects can be divided into two groups: one consists of non–dose related effects such as pancreatitis, fever, malaise, rash, nausea, diarrhoea and hepatitis; the second consists of dose-related effect, such as leucopoenia, thrombocytopenia, infection, malignancy and

hepatitis.[11] Thomsen et al. showed that during AZA therapy 5/439 children with acute lymphoblastic leukaemia developed secondary myelodysplasia or acute myeloid leukaemia.[12] More recently, three case reports have described new side effects of AZA. In the first case report, de Boer described a 35-year-old man who presented with hepatitis, nodular regenerative hyperplasia, veno-occlusive disease, fibrosis and sinusoidal dilatation upon treatment with AZA. Hepatitis is considered to be an idiosyncratic reaction to AZA, the other side effects are considered to be signs of dose-dependent hepatotoxicity.[13] The second case showed recurrent atrial fibrillation in a patient with ulcerative colitis treated with AZA. Atrial fibrillation can be a serious but reversible adverse effect.[14] The third case report showed that immunosuppressive drugs such as corticosteroids or AZA could induce hepatitis B virus reactivation in patients carrying hepatitis B virus surface antigen, albeit much less frequently than chemotherapy treatments.[15]

2. THE MAJOR PATHWAYS FOR AZA METABOLISM. INFLUENCE OF GENETIC VARIATIONS IN RESPONSE TO THE DRUG

AZA and 6-MP undergo extensive first-pass metabolism, and the contribution of the three major metabolic pathways (TPMT, XO and HGPRT) to individual variations in its metabolism has been studied by pharmacogenetics.[16]

2.1. Thiopurine S-methyltransferase

Thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme that catalyzes the S-methylation of aromatic and heterocyclic sulfhydryl compounds, including cytotoxic and immunosuppressive thiopurine medications such as mercaptopurine, azathioprine and thioguanine.[17] While TPMT is an important determinant of toxicity with these medications, no endogenous substrate is known for this enzyme and its biological role remains obscure.[18]

The *TPMT* gene is located on chromosome 6 (6p22) and consists of 10 exons and 9 introns (**Figure 3**). It encodes a 245 amino acid protein with a predicted molecular mass of 35 kDa. A processed pseudo gene of *TPMT* has also been described with 96% homology to the *TPMT* gene; it is located on chromosome 18q21.1.[10]

The 5' promoter region has a 71% GC content; it has no consensus sequences for TATA box or CCAAT elements, and contains several binding sites for well-characterized

transcription factors including Sp1, NF-k, AP-2 and KROX-24.[18] Variable number tandem repeats have been identified in the 5'UTR.[19]

TPMT enzymatic activity is inherited as an autosomal codominant trait. In humans, it is controlled by a genetic polymorphism at the *TPMT* locus and approximately 1 in 300 individuals has very low TPMT activity, 11% have intermediate activity, and 89% show normal/high activity.[20] Several clinical studies found that high methylators may be under-treated with conventional doses of thiopurine drugs, whereas intermediate and deficient methylators are recognized to be at risk of moderate to profound haematopoietic toxicity when treated with standard doses of the medications. These observations are explained by an inverse relationship between TPMT activity and the production of TGNs.[21]

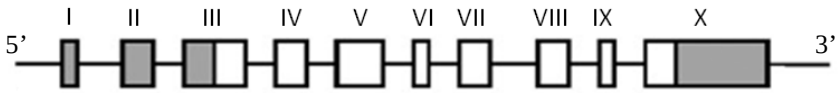


Figure 3. Gene structure of thiopurine S-methyltransferase (*TPMT*). Solid white boxes, coding exonic sequence. Solid grey boxes, non-coding exonic sequence. Sizes of the exons, but not the introns, are shown proportional to their real lengths.

2.2 *TPMT* allelic mutations

Many rare genetic variants exist in the human population, but most of the heterozygosity in the population is attributable to common alleles, that is those that are present in the general population at a frequency of >1% (single nucleotide polymorphisms). Single nucleotide polymorphisms (SNPs) can occur in the coding sequence or in the intronic region. If they are in the coding sequence, they could or could not determine an amino acid change in the protein sequence. If they are in the intronic sequence, they could modify mRNA splicing, causing the synthesis of a mutated protein.

Currently, a total of 35 different *TPMT* allelic variants have been characterized. They may be associated with a decreased TPMT activity and/or induced toxicity. Inactive *TPMT* alleles contain mutation(s) in the open reading frame (ORF), resulting in the synthesis of a protein that is more rapidly degradable, or in intron-exon junction, leading to an altered mRNA splicing. Some SNPs in *TPMT* gene cause a structural disruption and misfolding of the clinically important drug-metabolizing enzyme. Misfolded proteins can be removed

from the cell by degradation, but they can also form aggregates. Aggregosome formation represents a unique process by which cells can remove misfolded TPMT protein.[22]

The "wild type" *TPMT* allele, which determines high enzyme activity, is defined as *TPMT**1. The mutant alleles can be divided in four groups: *TPMT* alleles associated with normal activity; *TPMT* alleles suspected to decrease enzyme activity; inactive *TPMT* alleles and intronic *TPMT* alleles.

The first group includes those polymorphisms which do not cause alteration in the TPMT enzymatic function. Two silent polymorphisms (no amino acid change) were identified in the fifth (c.C339T, p.Thr113Thr) and the seventh (c.C474T, p.Ile158Ile) exons; both alleles are termed *TPMT**1S.[23] Other polymorphisms (C-178T, T-30A, G-23T) were detected in the 5'UTR region in exon 1, a non-coding region of the *TPMT* gene. c.C-178T is indicated as *TPMT**1A; the other two have not been named[24,25] (**Figure 4**).

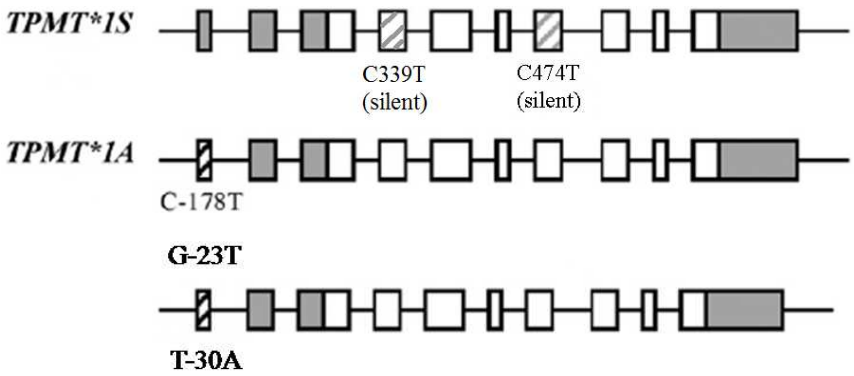


Figure 4. *TPMT* alleles associated with normal activity. Hatched boxes, exons containing at least one mutation; white boxes, exons with normal sequence. Solid gray boxes, non-coding exonic sequence. Sizes of the exons, but not the introns, are shown proportional to their real lengths.

The second group includes the alleles which present non-synonymous mutations resulting in amino acid change and are suspected to decrease TPMT activity, even if a clear and indisputable association is still lacking. These mutations are extremely rare. *TPMT**7 and *TPMT**10 are characterized by a non-synonymous mutation in exon 10 (c.T681G, p.His227Glu) and in exon 7 (c.G430C, p.Gly144Arg), respectively. They were identified in patients treated with conventional doses of AZA who developed severe leucopenia. *In vitro* experiments showed that the *TPMT**7 and *TPMT**10 allelic variants are a loss-of-function, but the molecular/structural basis for the detrimental influence of the amino acid change on TPMT activity remains to be elucidated.[26] *TPMT**8, which contains a single nucleotide transition (c.G644A) leading to an amino acid change at codon 215 (p.Arg215His), was found in an Afro-American cohort of patients with intermediate

activity. This mutation was not found in any other Afro-American individuals and the mechanism by which it is associated with low TPMT activity is still unclear.[27] Schaeffeler et al. found *TPMT**9 (c.A356C, p.Lys119Thr) in two individuals, one with an intermediate activity and the other with a normal TPMT activity, so this polymorphism does not seem to be associated with low TPMT activity. *TPMT**17 (c. C124G, p.Gln42Glu) and *TPMT**18 (c.G211A, p.Gly71Arg) were each described in only one patient. They seem to be associated with decreased TPMT activity leading to an intermediate TPMT phenotype.[28] These mutations were not found in other individuals tested and the mechanism by which they are associated with low TPMT activity is unknown. Hamdan-Khalil et al. during a routine genotyping of patients with Crohn's disease identified one non-synonymous mutation, *TPMT**19 (c.A365C, p.Lys122Thr), in a Caucasian patient who had an apparent normal TPMT activity.[21] *TPMT**20 consists of a mutation in exon 3 (c.G106A, p.Gly36Ser), with a frequency of 0.003 in the Japanese population. The Gly36Ser substitution is closely located to the methyl donor S-adenosylmethionine binding site of the N-terminal end of TPMT. Thus, the mutation is expected to alter the catalytic properties of TPMT, but this has yet to be demonstrated.[29] *TPMT**20* (so called thereafter to distinguish it from TPMT*20 previously identified in the Japanese population) and *TPMT**22 consist in an amino acid change in exon 10 (c.A712G, p.Lys238Glu) and in exon 7 (c.G488C, p.Arg163Pro), respectively.[30] *TPMT*20** and TPMT**22* alleles appear to be extremely rare in the Caucasian population, and *TPMT**20* is associated with intermediate activity, albeit at levels close to normal activity.[30] *TPMT**23 is a rare polymorphism which consists in a substitution in exon 8 (c.C500G) that determines an amino acid change (p.Ala167Gly). Alignment of human *TPMT* and orthologs indicated that Ala is highly conserved at this position, and the substitution of Ala for Gly at position 167 was predicted to be deleterious, resulting in defective enzyme function.[31] *TPMT**24 is a rare non-synonymous mutation located in exon 8 c.T634C that results in p.Gln179His substitution. The definite functional status of *TPMT**24 remains difficult to ascertain and further *in vitro* analyses are needed to demonstrate its impact on TPMT expression and/or stability[32] (**Figure 5**).

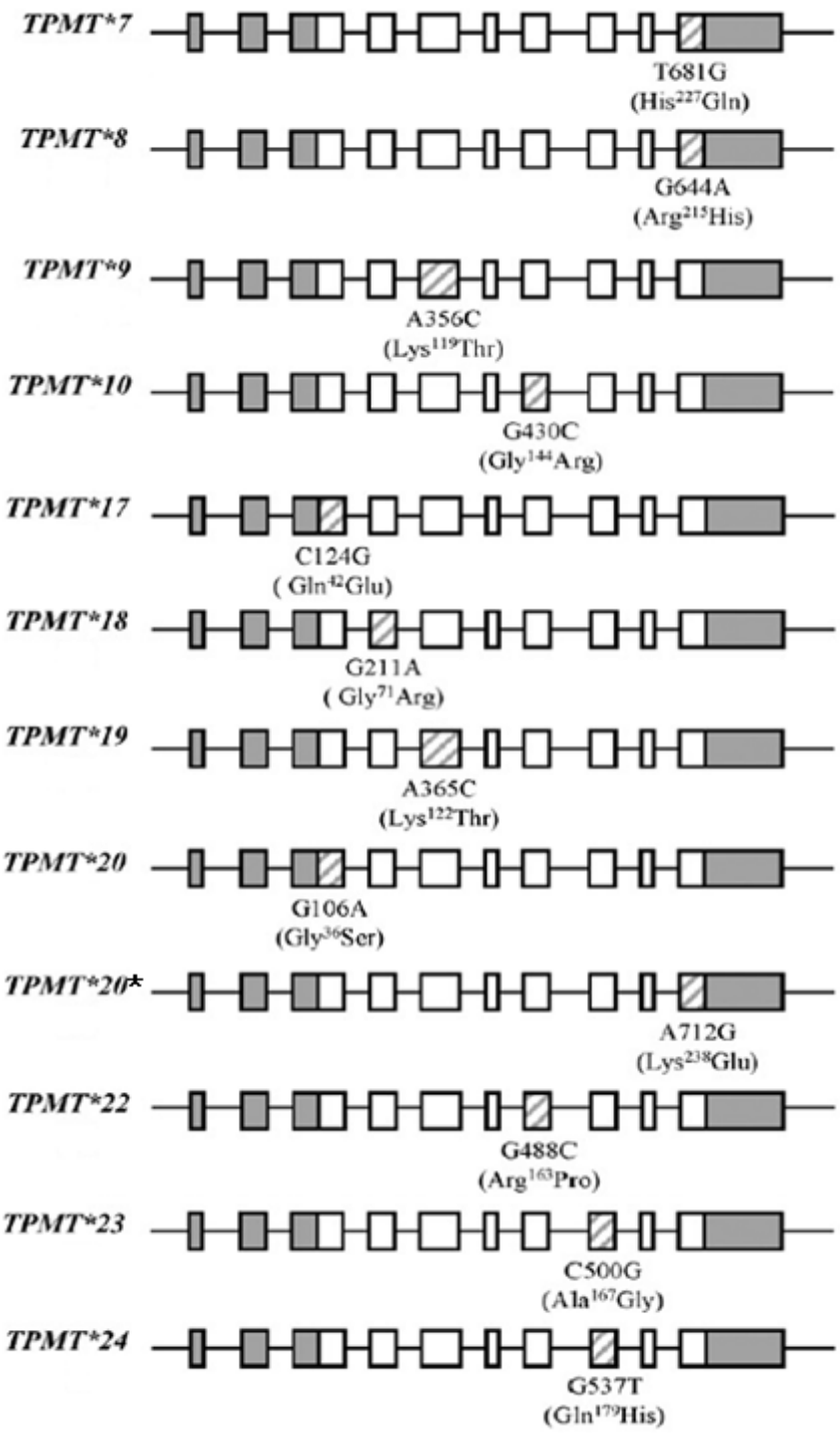


Figure 5. *TPMT* alleles suspected to decrease activity. Hatched boxes, exons containing at least one mutation; white boxes, exons with normal sequence. Solid gray boxes, non-coding exonic sequence. Sizes of the exons, but not the introns, are shown proportional to their real lengths.

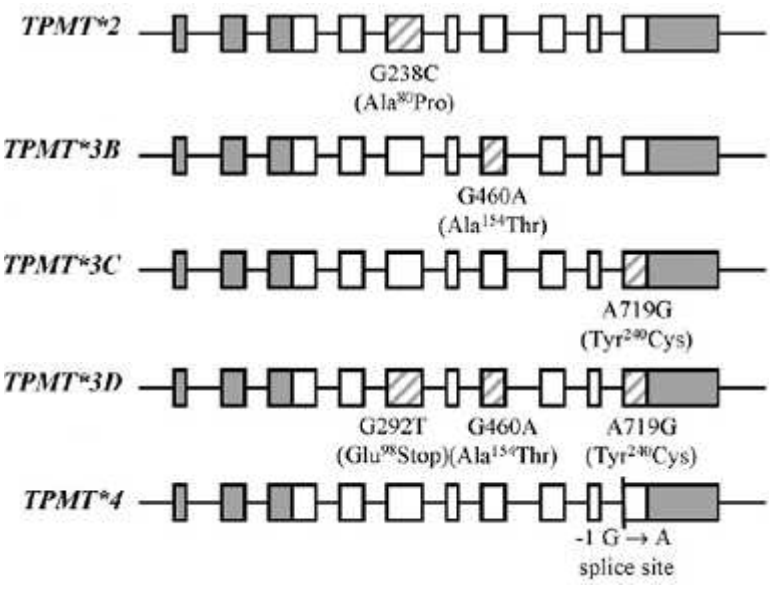
The third group includes polymorphisms of *TPMT* gene which cause an amino acid change responsible for the inactivity of the protein. *TPMT*3A*, *TPMT*3B TPMT*3C* and *TPMT*2* are the most prevalent mutant alleles in Caucasians, Asians, and Afro-Americans, comprising approximately 95% of *TPMT* mutant alleles in these populations. *TPMT*2*

contains a single nucleotide change in exon 5 (c.G238C, p.Ala80Pro), *TPMT*3B* includes a single mutation in exon 7 (c.G460A, p.Ala154Thr), *TPMT*3C* contains a single mutation in exon 10 (c.A719G, p.Tyr240Cys) and *TPMT*3A* contains two mutations, one in exon 7 (c.G460A, p.Ala154Thr) and the other in exon 10 (c.A719G, p.Tyr240Cys). Each of these mutant alleles has been shown to cause TPMT deficiency. *TPMT*3A* is the most prevalent mutant allele in Caucasians, whereas *TPMT*3C* is the predominant *TPMT* mutant allele in Asian, African, and Afro-American populations.[33,34]

*TPMT**2, assessed in a yeast heterologous expression system, led to a 100-fold reduction in TPMT catalytic activity compared to the wild-type, despite a comparable level of mRNA expression.[35] *TPMT**3A is rapidly degraded by an ubiquitin-proteasome mediated process, with the involvement of molecular chaperones.[36] This observation is in accordance with a growing body of evidence demonstrating that an alteration in protein level is a common mechanism underlying the functional effects of genetic polymorphisms involving non synonymous SNPs, most often as a result of accelerated protein degradation.[37] In contrast, *TPMT**5 (c.T146C, p.Leu49Ser) might alter the enzymatic function of the protein.[34]

*TPMT**3D is a haplotype which consists in the *TPMT**3A mutations plus another polymorphism at nucleotide 292 (c.G292T) in exon 5, which replaces the codon for Glu98 by a frame-stop codon (p.Glu98X). It determines intermediate enzyme activity. *TPMT**4 has a G-1A transition at the intron 9/exon 10 splice junction that modifies the final nucleotide of the intron at the invariant 3'-acceptor splice site sequence required for RNA processing. *TPMT**6 is characterized by a substitution A>T at nucleotide 539 (c.A539T) in exon 8 resulting in a p.Tyr180Phe substitution. This allele was identified in a Korean subject and caused low TPMT activity.[38] Schaeffeler et al. identified a mutation, c.G395A, in exon 6, that results in p.Cys132Tyr, which defines allele *TPMT**11; the molecular mechanism for loss of catalytic activity has not been elucidated. However, similarly to *TPMT**2 and *3A, a post-transcriptional mechanism with enhanced proteolysis of the mutant protein seems to be the most likely explanation.[39] *TPMT**12 (c.C374T, p.Ser125Leu) results in enzyme with significantly reduced intrinsic clearance values, about 3.5-fold lower than wild-type TPMT. Based on *in vitro* data, Hamdan-Khalil et al. postulated that *TPMT**12 and another demonstrated allele, *TPMT**13 (c.A83T, p.Glu28Val), correspond to loss-of-function alleles of TPMT.[26] *TPMT**14 is defined by a mutation in the first nucleotide of the coding sequence (c.A1G, p.Met1Val). According to the NCBI ORF finder, the theoretical consequence of *TPMT**14 may be a truncated

protein: use of a downstream ATG as a start codon will result in a shortened protein of 170 amino acids compared to the wild-type 245 amino acid-long TPMT. Allele *TPMT**15 has a G-1A transition at the intron 7/exon 8 splice junction of the *TPMT* gene. One possible consequence is the deletion of the entire exon 8 in the final protein, resulting in a frame shift and a premature stop codon in exon 9 that would lead to a truncated 173 amino acid-long protein.[40] Hamdan-Khalil[21] genotyped two Crohn's disease patients by *TPMT* cDNA sequencing, revealing the presence of one non-synonymous mutation in exon 7 corresponding to c.G488A, p.Arg163His substitution (*TPMT**16). *TPMT**16 can be considered a very rare *TPMT* variant with a frequency <0.1%. This allele encodes a protein with a severely impaired catalytic efficiency and could then be considered as a non-functional allele in the context of phenotype prediction via genotyping tests.[21] *TPMT**21 consists in a mutation in exon 4 (c.C205G, p.Leu69Val) and leads to a substantial decrease in TPMT activity *in vivo*.[30] The *TPMT**25 (c.T634C, p.Cys212Arg) allele is not expected to occur frequently in Caucasian populations. The Cys212Arg mutation is located on the protein surface away from the catalytic site. This solvent-exposed mutation is characterized by a modification in ionic charge that could destabilize the local electrostatic balance in the native protein, likely resulting in a protein with a significantly reduced catalytic activity toward 6-TG methylation compared to the wild-type enzyme, as observed through structural analysis. Consequently, this allelic variant can be regarded as belonging to the non-functional alleles of *TPMT*.[32] In the last years, three other mutations were found to be associated with low TPMT activity: c.T622C, p.Phe208Leu in exon 9 (*TPMT**26); c.T319G, p.Tyr107Asp in exon 5 (*TPMT**27); c.T611C, p.Ile204Thr in exon 9 (*TPMT**28)[41-43] (**Figure 6**).



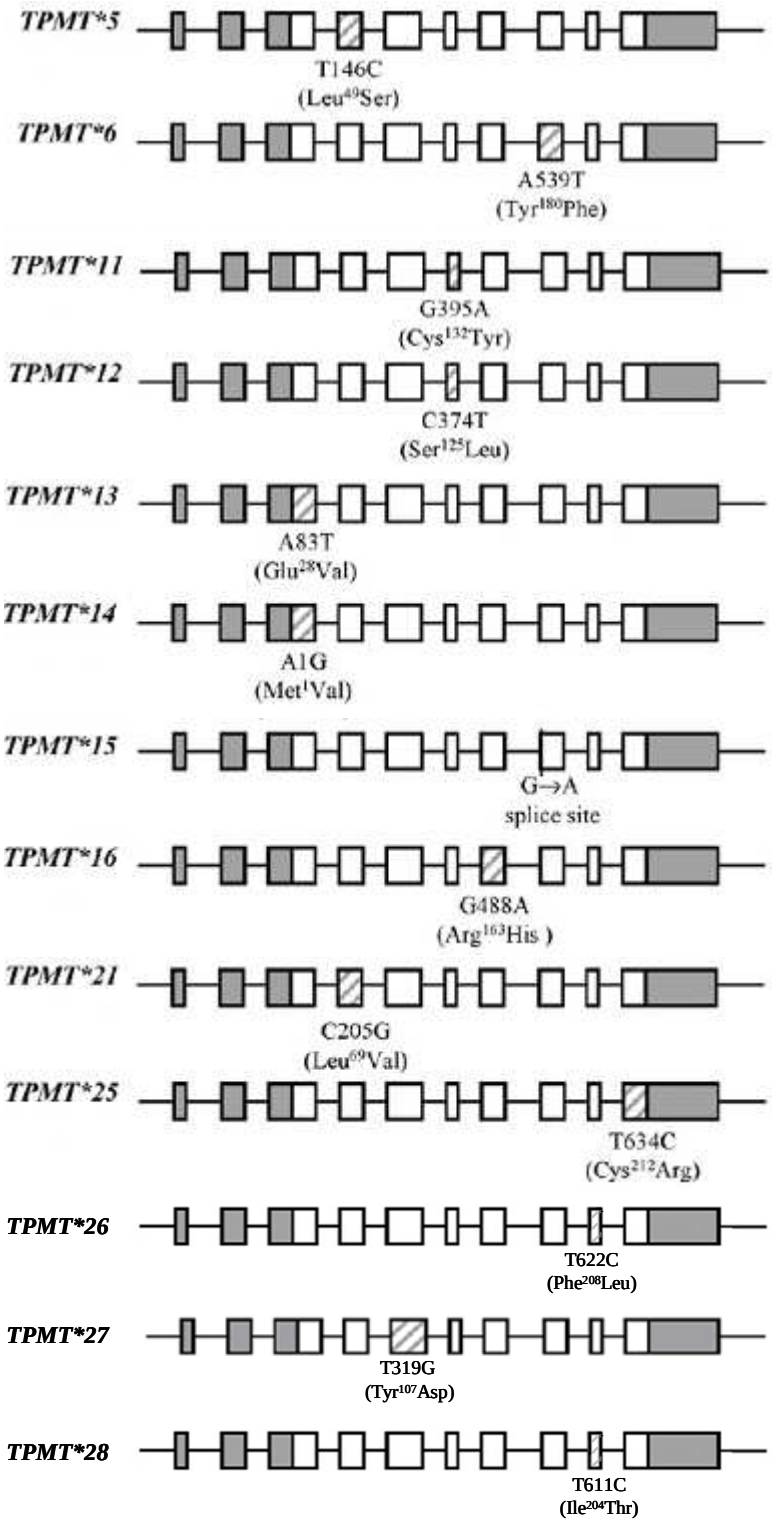


Figure 6. Defective *TPMT* alleles. Hatched boxes, exons containing at least one mutation; white boxes, exons with normal sequence. Solid gray boxes, non-coding exonic sequence. Sizes of the exons, but not the introns, are shown proportional to their real lengths.

Intronic polymorphisms were detected at nucleotide T140+114A and A141-101T in intron 3, and at nucleotide T366+58C in intron 5 (**Figure 7**). These polymorphisms do not cause

alteration in TPMT activity. However, definitive linkage analysis and the impact on enzyme activity must await characterization of a larger number of samples.[38,23]

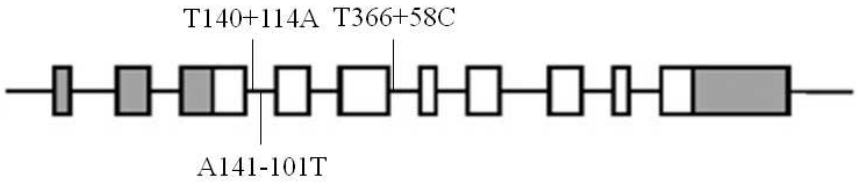


Figure 7. Intronic polymorphisms. Hatched boxes, exons containing at least one mutation; white boxes, exons with normal sequence. Solid gray boxes, non-coding exonic sequence. Sizes of the exons, but not the introns, are shown proportional to their real lengths.

2.3 Activity of TPMT

Neither the biological role of TPMT nor its endogenous substrate have been defined as yet, but we know that TPMT-deficient patients are at risk of severe and potentially fatal hematopoietic toxicity, unless their thiopurine dosage is decreased by 10-15 fold.[44,45] In addition to genetic polymorphisms, tissue-specific differences in *TPMT* gene expression have been documented, with the highest level in liver and kidney and relatively low levels in brain and lung. TPMT expression may also exhibit developmental regulation, as erythrocyte activity is higher (about 50%) in newborns than in healthy adults, resulting in a higher TPMT protein content in erythrocytes of newborns.[46] Erythrocyte TPMT activity is approximately 30% higher when children with leukaemia are taking mercaptopurine, and returns to pre-treatment levels when thiopurine therapy is stopped.[18]

2.4 Limitations in analysis of TPMT activity

Individuals who have inherited a mutated *TPMT* gene can develop a hematopoietic toxicity when treated with standard doses of 6-MP or AZA. The determination of TPMT activity in red blood cells could therefore be a predictive test for intolerance before starting with a thiopurine therapy. However, analysis of TPMT activity currently shows limits. Indeed, in particular cases, TPMT activity can not be determined correctly: when a patient heterozygous for TPMT SNPs has received blood transfusions from a wild-type individual (the test must then be made 60 days after blood transfusion); when the patient takes AZA already, because the presence of AZA induces a natural increase of TPMT activity by approximately 20% above the basal value; or when the patient takes other drugs such as

sulfasalazine or olsalazine, that are strong inhibitors of TPMT.[28] There is therefore a need for approaches to predict TPMT activity, and thereby drug response to AZA, that overcome the limitations of enzyme activity testing.

3. XANTHINE OXIDASE

Xanthine oxidase (XO) is a cytoplasmatic enzyme that catalyzes the last two steps of purine degradation: oxidation of hypoxanthine to xanthine and xanthine to uric acid with concomitant reduction of NAD^+ (dehydrogenase) or molecular oxygen (oxidase). XO oxidizes purines, pyridines, pterins, and aldehydes, and it is also involved in the metabolism of clinically significant drugs such as 6-MP, allopurinol (an XO inhibitor), pyrazinamide (an antituberculous agent),[47] and azathioprine. XO metabolizes approximately the 10% molar equivalent of a normal daily AZA dose, and it is excreted as urinary 6-tiouric acid (6TU).[48]

Xanthine oxidoreductase enzymes have been isolated from a wide range of organisms, from bacteria to man. All of these proteins have similar molecular weights and composition of redox centres.[49] XO and xanthine dehydrogenase (XDH) are both members of the molybdenum hydroxylase flavoprotein family and represent different forms of the same gene product. The two enzyme forms and their reactions are often referred to as xanthine oxidoreductase activity. XDH is a homodimer with a subunit molecular mass of approximately 150 kDa. Conversion of XDH into XO is a two-step process with the initial step involving the oxidation of the important protein thiol groups. Irreversible cleavage of a 20 kDa fragment from each of the subunit by a calcium-dependent protease subsequently converts XDH into XO.[50] The active form of the enzyme is a homodimer of 290 kDa molecular mass, with each of the monomers acting independently in catalysis. Each subunit contains one molybdopterin domain, two spectroscopically distinct [2Fe-2S] centres, and one FAD cofactor. The oxidation of xanthine takes place at the molybdopterin centre (Mo-pt) and the electrons thus introduced are rapidly distributed to the other centres by intramolecular electron transfer (**Figure 8**).[49] The highest XO activities are found in human liver and bowel as opposed to other tissues. High levels of XO have been associated with tissue injury in certain diseases and is considered to contribute to oxidative damage of cells through the generation of cytotoxic oxygen radicals, which are implicated as important pathological mediators in many clinical disorders, including ischemia-reperfusion injury, myocardial infarction, hypertension, and atherosclerosis.[51,52]

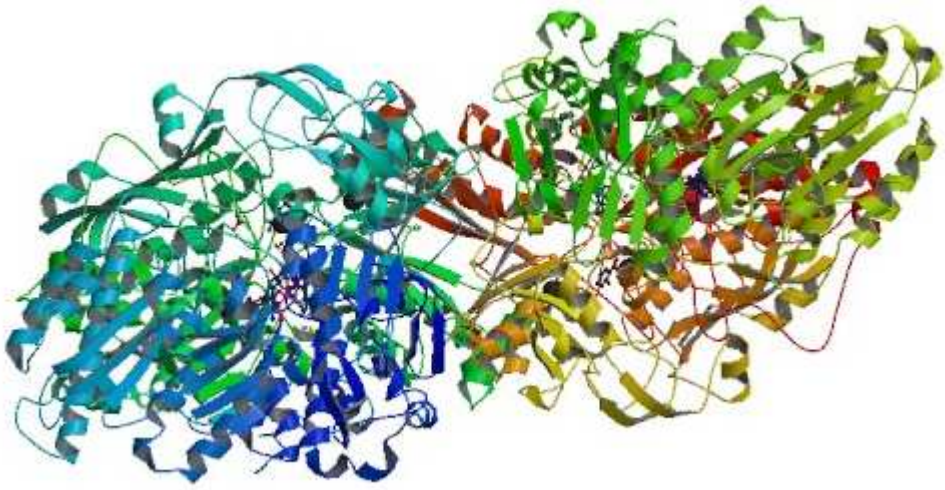


Figure 8. 3D structure of xanthine oxidase.

There is inter-individual variation in human liver XO activity, and about 20% of Caucasian subjects display relatively lower enzyme activity.[53] Hence, about 11% of Japanese[54] and 4% of Spanish individuals[55] are likely to be poor metabolizers of XO substrates and XO deficiency occurs in 2% of the general French population.[56]

XDH polymorphisms have been well characterized among Caucasians and Asians but not in Africans.[57] The human *XDH* gene has been mapped to chromosome 2p22. The *XDH* gene spans over 60000 bp of DNA and can be divided into 36 relatively small exons. Exon 36 is over 700 bp long and contains the last few translated amino acids and the entire 3'-untranslated region. Each of the remaining 35 exons consists of 50-280 bp. The 5' flanking region is characterized by a sequence similar to that of an initiation site, lacking the canonical TATA boxes and several binding sites for various transcription factors. This region appears to be that of a highly regulated TATA-less gene and this is consistent with its tissue specific expression and modulation by various stimuli at the transcriptional level.[50] The *XDH* gene encodes approximately 1330 amino acids and each monomer of XDH can be divided into three domains. The small N-terminal domain (residues 1 to 165, 20 kDa) contains both iron/sulphur cofactors and is connected to the second, FAD-binding, domain (residues 226 to 531, 40 kDa) by a long segment consisting of residues 166 to 225. The FAD domain is connected to the third domain by another linker segment (residues 532 to 589). The large third domain, residues 590 to 1335 (85 kDa), sequesters the Mo-pterin cofactor close to the interfaces of the Fe/S- and FAD-binding domains. The C-terminal domain spans residues 1310 to 1335[49] (**Figure 9**).

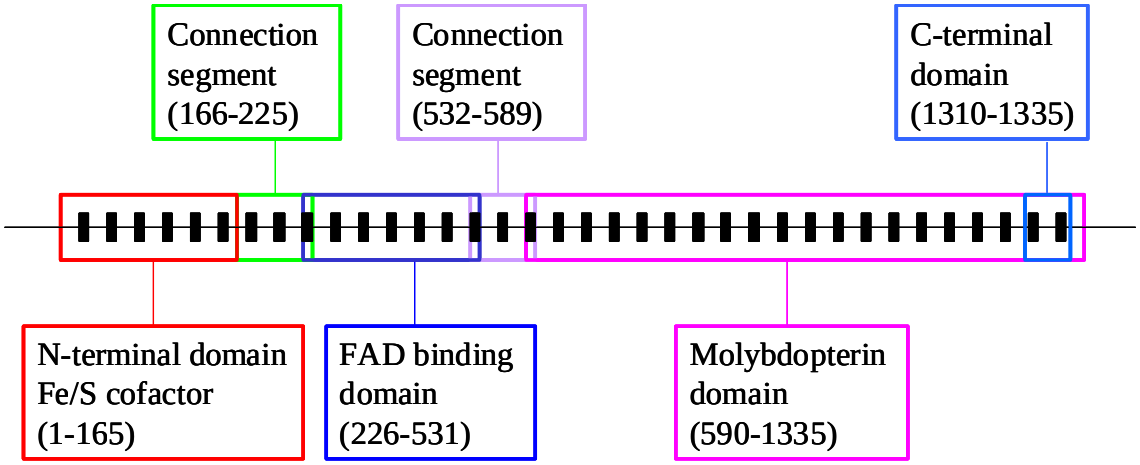


Figure 9. Scheme of *XDH* monomer and subdivision in functional domain. Exons are shown as black boxes. Note that the depiction of the *XDH* gene is not to scale.

The amino acids mainly involved in the active site in the iron-sulphur domain are cysteines implicated in disulphide bounds: Cys 43, Cys48, Cys51, Cys73, Cys113, Cys116, Cys148, Cys150; those implicated in the active site of the FAD binding domain are: Phe337, Asp360, Leu404, Lys422; and the amino acids involved in the active site of the molybdopterin domain are: Gln768, Phe799, Glu803, Arg881, Arg913, Phe915, Thr1011, Ala1080, Glu1262 (proton acceptor).[58]

3.1. *XDH* single nucleotide polymorphisms

Many mutations and SNPs are known to occur in the *XDH* gene. In 1997, Ichida et al. identified the first two *XDH* mutations in Japanese patients affected by xanthinuria. These two mutations determine deficiency in XO activity. One subject had a C to T base substitution at nucleotide 682 (c.C682T) that caused a CGA (Arg) to TGA (STOP) nonsense substitution at codon 228 (p.Arg228X). The duodenal mucosa from the subject had no xanthine dehydrogenase protein, despite normal mRNA levels. The other subject had a C deletion at nucleotide 2567 that generated a modified protein with a termination codon 72 amino acids after the mutation.[59] In 2000, Levartovsky et al. found in an Iranian-Jewish family a 1658insC insertion in exon 16 of the *XDH* gene, which causes a translation frame shift predicting a truncated protein of 569 amino acids with an altered C-terminal amino acid sequence. The predicted truncated protein is thought to be functionally compromised, since amino acid residues needed for the formation of the putative NAD

binding domain (encoded by exon 22), the molybdopterin binding site (encoded by exons 22 and 23) and the amino acid residues involved in the conversion of XDH to XO (exons 16, 27, and 36) are lost.[60] In 2001, Sakamoto et al. demonstrated a point mutation of C to T in nucleotide 445, which changes codon 149 from CGC (Arg) to TGC (Cys) (c.C445T, p.Arg149Cys), in one patient. The substitution of arginine by cysteine was identified near the cysteine residue that is proposed to participate in the iron-sulphur centre. Therefore, this substitution may change the tertiary structure of XO, leading to an absence of XO activity.[61] In 2003, Gok et al. identified a new mutation in exon 20, an A to T base change at nucleotide position 2164 (c.A2164T) indicating a non-synonymous substitution from AAG (Lys) to TAG (Tyr) at codon 722 (p.Lys722Tyr). The consequences on protein activity of this new mutation are unknown.[62]

Two large studies were conducted in 2008. Yang et al. genotyped the *XDH* gene of 48 subjects affected by hypertension. They identified three non-synonymous mutations (c.G514A, p.Gly172Arg; c.G2794A, p.Ala932Thr; c.A3326C, p.Asn1109Thr) and five synonymous variations (c.G627A, p.Glu209Glu; c.C837T, p.Val279Val; c.C2211T, p.Ile737Ile; c.T3030C, p.Phe1010Phe; c.G3717A, p.Glu1239Glu). Whether or not these mutations cause a modification in XO enzyme activity remains to be ascertained.[63] Kudo et al. genotyped the *XDH* gene of 96 unrelated Japanese volunteers. They found 19 non-synonymous and 3 synonymous mutations. They also studied the functional activity of XO in mutant subjects. The results showed a deficiency in enzyme activity in two variants (c.C445T, p.Arg149Cys and c.C2729A, p.Thr910Lys); low activity in six variants (c.C1663T, p.Pro555Ser; c.G1820A, p.Arg607Gln; c.C1868T, p.Thr623Ile; c.C2727A, p.Asn909Lys; c.C3449G, p.Pro1150Arg and c.G3953A, p.Cys1318Tyr); and hyper-activity in two variants (c.A2107G, p.Ile703Val and c.A3662G, p.His1221Arg).[47] Finally, in 2010, Jureca et al. found, in a patient affected by rheumatoid arthritis, a new deletion in exon 8 (c.641delC, p.Pro214QfsX4) resulting in a frame-shift change, with a change of proline into glutamine at residue 214. The reading frame stops four codons after the mutation and the resulting mutated protein has only 218 amino acids. This protein has the first small N-terminal domain (residue 3–164 amino acids), which contains the Fe-S cofactors, and an incomplete second, FAD-binding, domain (192–527 amino acids).[64]

4. PHARMACOGENETICS, PHARMACOGENOMICS AND PERSONALIZED MEDICINE

Adverse drug reaction in patients causes more than 2 million hospitalizations including 100000 deaths per year in the United States.[65] This adverse drug reaction could be due to multiple factors such as disease determinants, environmental and genetic factors. In order to improve the efficacy and safety of drugs and understand their clinical consequences, two rapidly developing fields – pharmacogenetics (focused on one or a few genes) and pharmacogenomics (focused on drug metabolism pathways) – are helping in genetic personalization of drug response. Indeed, many drug responses appear to be genetically determined and the relationship between genotype and drug response may have a very valuable diagnostic value. Identification and characterization of a large number of genetic polymorphisms (biomarkers) in drug-metabolizing enzymes and drug transporters in an ethnically diverse group of individuals may provide substantial knowledge about the mechanisms of inter-individual differences in drug response. Although individualized medications remain a challenge for the future, the pharmacogenetic approach to understand drug response should be further developed in order to identify biomarkers at the genetic level and apply this knowledge, not only to improve public health, thereby avoiding adverse drug reactions (by knowing in advance who should be treated with what drug and how), but also towards drug development[66] (**Figure 10**).

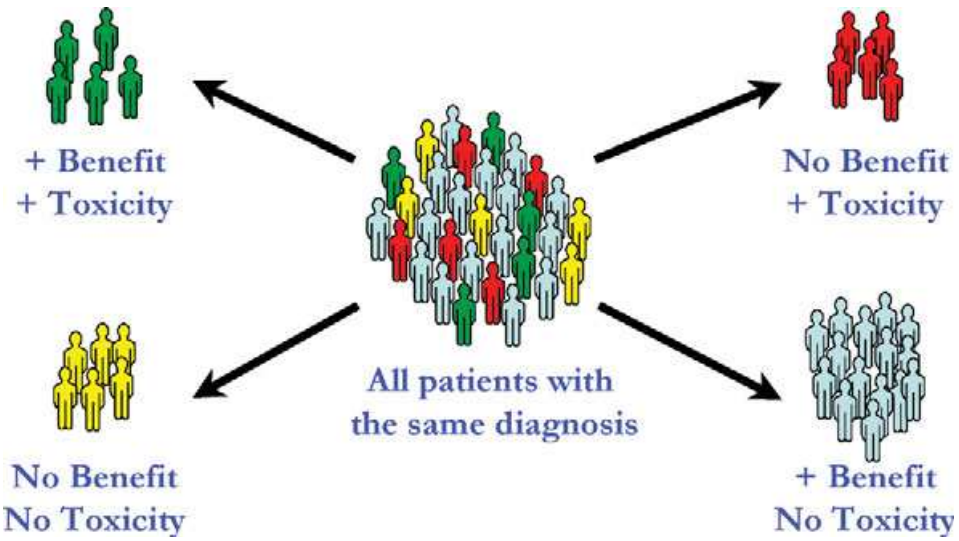


Figure 10. Pharmacogenomics and pharmacogenetics aim to identify patients at risk for toxicity or reduced response to therapy prior to medication selection.

Many genes encoding proteins involved in the metabolism, transport, and mechanism of action of medications are known to exhibit polymorphism in humans, but the use of this knowledge in routine clinical practice is limited. In fact, even if a gene has a large effect on a drug's pharmacokinetics or pharmacodynamics, the presence of a single-nucleotide polymorphism in that gene will not provide an unequivocal answer but, rather, will indicate the likelihood that an individual patient will show an altered drug response.[67] Recent pharmacogenetic studies have mainly focused on immunosuppressive agents including corticosteroids, AZA, metotrexate and infliximab.[68] Variability in the expression of pharmacokinetic (PK) genes can either slow or accelerate the uptake, conversion, or excretion of drugs, which can result in compounds or their active metabolites being eliminated before achieving a therapeutic effect, or remaining too long or in too high concentration so that the risk of adverse events increases.[69] An example of a gene encoding an enzyme involved in PK process is TPMT. TPMT catalyzes the S-methylation (inactivation) of the thiopurine drugs, mercaptopurine, AZA, and thioguanine. In hematopoietic tissue, TPMT is the predominant inactivation pathway, so patients who inherit TPMT deficiency accumulate excessive levels of active thioguanine nucleotides after receiving standard doses. Numerous studies have shown that TPMT-deficient patients are at high risk of severe, sometimes fatal, hematologic toxicity; TPMT heterozygotes have an intermediate risk of hematologic toxicity. Actually more than 98% concordance exists between *TPMT* genotype and phenotype, and genotyping is highly sensitive (90%) and specific (99%) in identifying patients who have inherited one or two non-functional TPMT alleles.[67]

The impact of common polymorphisms in *TPMT* on clinical outcome with AZA treatment has become a classic example of the application of pharmacogenetics, and is one of the first development in this field to be widely adopted in clinical practice. The other gene that has been subjected to significant enquiry in the context of AZA treatment is XO. The bioavailability of 6-MP has been estimated as approximately 5-37% owing to the presence of large amounts of XO in the liver. This bioavailability increases when XO inhibitors such as metotrexate are administrated concomitantly with 6-MP. Competitive inhibition of XO with allopurinol pre-treatment results in 5-fold increase in the bioavailability of oral 6-MP. Concurrent use of allopurinol and 6-MP has been reported to cause severe myelotoxicity. 6-MP-induced adverse effects may increase in poor XO metabolizer.[70] Wong reported that the levels of active metabolites are undetectable when XO activity is strong, in fact in a

patient with high XO enzyme activity, 6-TGNs could not be detected despite AZA administration of 150 mg/die.[71]

AIM OF THE STUDY

Single nucleotide polymorphisms influence the response to thiopurine drugs. Many defective and potentially deficient alleles were identified for *TPMT* that cause intolerance to azathioprine. *XDH* gene also showed many different polymorphisms that determine defective or hyper-XO activity, which have not yet been tested for possible association with AZA intolerance or unresponsiveness.

Currently, the pharmacogenetics knowledge does not fully to explain the unresponsiveness and the majority of intolerant patients to azathioprine, and there is a need to further define pharmacogenetic profiles that can guide pharmacological therapy.

Since azathioprine is used in the treatment of different diseases, the aim of our study is to perform genotypic analysis of *TPMT* and *XDH* genes, involved in the catabolism of AZA, to identify those new polymorphisms and possibly novel haplotypes that could determine adverse response or unresponsiveness to azathioprine.

MATERIALS AND METHODS

1. SUBJECTS

DNA samples from a cohort of 71 Italian patients with autoimmune diseases treated with AZA were obtained from the DNA bank of the Neurology IV Unit at Neurological Institute Carlo Besta, Milan, Italy or extracted from peripheral blood. Written informed consent for DNA storage and use for research purposes was obtained from all patients, as required by the ethical committee of Foundation Neurological Institute "Carlo Besta", which approved the study. The patients underwent AZA treatment and were grouped according to their response to the appropriate dose of AZA (2-3 mg/kg per day): responsive after at least one-year follow-up if they achieved pharmacological remission, or if clinical course was stable and steroid dosage had been reduced to at least 50% of pre-AZA level; intolerant, if they experienced persistent side effects upon AZA treatment, such as described by Ansari et al.,[72] which could only be abrogated by withdrawal of the drug; or unresponsive if they showed neither clinical improvement nor adverse effects after a one-year treatment[73]. Clinical data were obtained retrospectively from the records available at the Neurology IV Unit database and reviewed by a doctor blinded for the genotyping results.

2. DNA EXTRACTION

Genomic DNA was extracted according to standard procedures from fresh or frozen samples of whole peripheral blood from AZAtreated patients and healthy individuals and stored at -20°C pending assays.

3. PCR AND DNA SEQUENCING

The genomic structure and intron boundary sequences were derived from the NCBI's RefSeq database, access numbers NG_012137.1 (genomic sequence) and CCDS4543.1 (coding sequence) for *TPMT* gene and access numbers NG_008871.1 (genomic sequence) and CCDS1775.1 (coding sequence) for *XDH* gene. Primers specific for *TPMT* exons 3, 4, 6, 8, and 9, and/or intronic sequences containing known polymorphisms[38] were identified using the Primer Express software package (Applied Biosystem, Foster City, CA); primers

specific for exons 5, 7, and 10 of *TPMT* gene and *XDH* gene were previously reported.[23,47] PCR amplification was carried out in a total volume of 25 µl containing 200 ng of genomic DNA, 1x DyNAzyme Buffer (Finnzymes, Celbio, Rho, Italy), 0.2 mM each GeneAmp dNTP (Applied Biosystems), 0.04 U DynaZyme DNA I DNA Polymerase (Finnzymes), and 0.5 µM of the specific primer pair. PCR reactions were performed on a GeneAmp PCR System 2700 (Applied Biosystem). Thermal cycling conditions to amplify exons 5, 7, and 10 of *TPMT* gene and all exons of *XDH* gene were according to Schaeffeler et al. and Kudo et al.[23,47] For the other *TPMT* exons, thermal cycling conditions were as follows: 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C, followed by 7 min at 72°C. PCR products (10 µl) were analyzed on 2% ethidium bromide-stained agarose gel electrophoresis, purified (2.5 µl of each) using ExoSAP-IT (GE Healthcare, Milan, Italy), and sequenced using Big Dye terminator v3.1 Cycle Sequencing Kit and a ABI 3100 Genetic Analyzer (Applied Biosystem). The sequences were analyzed using SeqScape v2.1.1 (all from Applied Biosystem). New nucleotidic changes identified were confirmed by at least two different PCR and sequencing reactions. The known *TPMT* polymorphisms that were analyzed in this study are listed in **Table 1** and those of *XDH* gene in **Table 2**.

The new nucleotidic changes that were detected in this study in our cohort of 71 patients were classified as mutations if they occurred with a frequency < 1% and not found in any of the control individuals; they were classified as polymorphisms if they occurred with a frequency >1%, including in control individuals.

Table 1. *TPMT* polymorphisms analyzed by DNA sequencing in this study

Exon	SNP (dbSNP identifier where available) [a]
3	T-30A[b]; G-23T[b]; A1G (rs9333569); T30C[c]; G42A[c] (rs111674665); C70T[c] (rs16880307); A83T (rs72552742); T97G (rs72552741); G106A; C124G
4	T146C (rs72552740); C205G (rs79321208); C211G
5	G238C (rs1800462); C244T (rs111901354); G292T (rs72552739); C339T[c] (rs17839843); G340A (rs115106679); A356C; A365C
6	C374T; G395A (rs72552738)
7	A423C[c] (rs16880276); A424C; G430C (rs72552737); G460A (rs1800460); C474T[c] (rs2842934); C487T (rs112339338); G488C; G488A
8	C500G (rs74423290); G537T (rs6921259); A539T (rs75543815); T547C (rs113525437)
9	T611C (rs79901429); T622C (rs72556347)
10	T634C; G644A (rs56161402); T646C (rs17849975); T681G (rs72552736); A712C; A719G (rs1142345)
Intron	
3	T140+114A (rs3931660); A141-101T (rs12529220)
5	T366+58C (rs2518463)
7	G496-1A
9	A627-1G

[a]www.ncbi.nlm.nih.gov/snp; [b]Polymorphism located upstream of the start codon; [c]Silent polymorphism

Table 2. *XDH* polymorphisms analyzed by DNA sequencing in this study

Exon	SNP (dbSNP identifier where available) [a]
1	G7A (rs144150776);G18T (rs113323020)
2	T78C (rs112608986)[b]; A98G(rs146551089)
3	G104A (rs139321199); G135A (rs150847028)[b]; C178T (rs184424277); G192T (rs182317768)
5	G328A (rs147286954); C339T (rs118042822)[b]; C385T (rs145064943); G397A (rs45447191); C405T (rs72549370)[b]; C426G (rs192763480)[b]
6	C445T; G446A (rs148108999); C463A (rs145413551); C468T (rs146749952)[b]
7	500delG (rs45462996); 501delT (rs143464813); T501G (rs143464813)[b]; G514A (rs45523133); C561G (rs191822672)
8	T572A (rs1126469); G627A (rs45575032)[b]; T632A (rs147724051); T636C (rs140225517)[b]; 641delC
9	G671A (rs112466737); C682T; C697T (rs149487595); A704T (rs45469499); C762T (rs140858800)[b]; G763A (rs151251011)
10	G822A (rs113345553)[b]; T825C (rs145596057)[b]; A829G (rs149853739); G835A (rs150515123); C837T (rs4407290)b; C859G (rs138674014); T876A (rs149541044); C882T (rs138246330)[b];G883A (rs142335345)
11	C981T (rs147062442)[b]; C1037T (rs142388231)
12	G1042A (rs192453952); C1101T (rs149225690)[b]; T1122A (rs1126470)[b]
13	C1134T (rs45612738)[b]; C1170G (rs146169510); A1184T (rs34929837); C1199T (rs140879910); G1238T (rs148024324)
14	C1274G (rs138649664); G1329A (rs45471294)[b]; G1357A (rs185036611)
15	G1495A (rs150262349); C1509G (rs45612839)[b]; T1514C (rs140685462); G1523A (rs148603901); C1535G (rs74874345); G1554A (rs143879399)[b]; C1565A (rs139199704); G1581A (rs138859065)[b]; G1597A (rs147476468); C1602T (rs61731080)[b]
16	G1617T (rs117399514)[b]; A1624G (rs149150797); C1629T (rs140651875)[b]; G1653A (rs80188867)[b]; 1658InsC; C1663T (rs45577338); G1677T (rs183162063)
17	T1707G (rs144540305)[b]; G1724A (rs147195351); C1726T (rs139772558); G1749A

	(rs141831710)[b]; A1751C (rs45491693); C1773T (rs145108969); G1774A (rs141314828); C1785T (rs140066757)[b]; C1800T (rs139964785)[b]; G1801A (rs61731081); G1820A (rs45442092); G1851T (rs45442398)
18	C1868T (rs45448694); G1885A (rs148464316); T1896G (rs14959399); C1911T (rs142197675)[b]; A1936G (rs17323225); C1970T (rs143481133)
19	C2022T (rs185713366)[b]; A2037C (rs139998082)[b]; A2069G (rs181111808); C2084T (rs142001530)
20	T2106A/G (rs72549368)[b]; A2107G (rs17011368); T2108A I (rs142848703); A2126G (rs146345594); T2159C (rs72552333); A2164T (rs72549367); G2176A (rs61731083)
21	T2206A (rs141470276); 2206delT (rs141470276); T2210C (rs72549366); C2211T (rs2295475)[b]
22	G2340T (rs150879057); G2347A (rs138717520); C2350T (rs141938849); C2359T (rs148904866); G2360A (rs149717617); A2362T (rs141050887); G2372A (rs143887178); C2413T (rs140007233); G2437C (rs142951412); C2441T (rs148585342); T2452C (rs139272198)
23	C2460T (rs143539472); G2477A (rs150297607)
24	2567DelC; G2581T (rs76810353); G2602C (rs148269820); C2626G (rs140982129)
25	T2634C (rs146994573)[b]; A2660G (rs145291309); T2684A (rs139272445); C2727A (rs566362); C2729A (rs669884); C2729T (rs669884); C2772T (rs143471207)[b]; G2794A
26	C2827T (rs144514609); G2828A (rs138396004); G2851A (rs142675390); T2858G (rs137975977); G2877A (rs151059034)[b]; G2911A (rs148921536)
27	T3018C (rs144883930)[b]; C3027T (rs72549365)[b]; T3030C (rs1884725)[b]
28	C3084T (rs45604135)[b]; C3126T (rs141214032)[b]; A3133G (rs142402298)
29	G3203A (rs186225688); A3205G (rs146283435); G3234A (rs111951522)[b]; C3249T (rs45562835)[b]; G3271C (rs45619033)
30	A3326C (rs45547640); C3348T (rs150419940)[b]
31	A3375G (rs45592238)[b]
32	C3449G (rs1042036); T3488C (rs142988357)
33	C3526T (rs45624433); C3530T (rs148972840); T3536C (rs139515054)
34	C3647A (rs143981573); G3660A (rs146264603)[b]; A3662G; C3665T (rs148235835); C3667T (rs116290580); G3717A (rs207440)[b]; G3724A (rs183714440); C3736T (rs142329784); G3737A (rs140606023); C3738T (rs151338723)[b]; A3764T (rs147253350); G3771A (rs45594136)[b]
35	C3792T (rs113929916)[b]; G3841A (rs181743036); T3864C (rs148477626)[b]; A3875G (rs73922346); C3886T (rs45564939); G3887A (rs139691722); G3909A (rs143920171)[b]; G3920A (rs138249576); C3930T (rs138936101); A3937G (rs141335716)
36	G3953A (rs2295474)
Intron	
2	C100+27T (rs206801)
4	A307-33G (rs10175754)
17	G1857-4A (rs17395175)
20	G2197+42C (rs761926); G2197+68A (rs2281550)
23	A2544+35T (rs2365842)
25	G2824-35A (rs13419410)
27	C3052-12G (rs13415401)
30	A3352-30C (rs207454)
34	C3774+32G (rs45591431)
3’UTR	A4002+518G (rs1042039); A4002+1477G (rs6710015); C4002+1516T (rs1054889); G4002+1555T (rs6752058)

[a]www.ncbi.nlm.nih.gov/snp; [b]Silent polymorphism

4. SPLICE SITE PREDICTION

Splice site of *TPMT* intron was predicted through BDGP: *Splice Site Prediction* by Neural Network (http://www.fruitfly.org/seq_tools/splice.html), which runs the NNSPLICE 0.9 version of the splice site predictor.

5. *XDH* NEW EXONIC MUTATION ANALYSIS

To evaluate the effect of exonic mutations on protein structure and enzymatic activity, we calculated the protein stability (ddG) upon mutation(s) using FoldX forcefields algorithm.[74] The default FoldX parameters were used (Temperature: 298 K; Ion strength: 0.05 M; pH: 7; Van der Waals Design: 2): if the mutation destabilizes the structure, ddG is increased, whereas stabilizing mutations decrease the ddG. Since the FoldX error margin is around 0.5 kcal/mol, changes in this range are considered insignificant. The XO crystallography structure (pdbID: 2E1Q) [http://nist.rcsb.org/pdb/explore/explore.do?structureId=2E1Q] was used as template. First, the tertiary structure was energetically minimized using the "RepairPDB" function of foldX.[75]

6. STATISTICAL ANALYSIS OF *TPMT* AND *XDH* GENES

Hardy-Weinberg equilibrium (HWE, p-value >0.05) and minor allele frequency (MAF <1%) were tested for each locus. Calculation and visualization of pairwise linkage disequilibrium (LD) between SNPs were carried out using the software Haploview version 4.1. Pairwise LD among the SNPs was assessed using standard coefficient D' and LD coefficient expressed as r^2.[76] The degree of LD among SNPs to partition haplotype blocks was examined using the Solid Spine algorithm in Haploview v4.1. Haplotype-trait associations were tested within the 18 kb block region of *TPMT* and 71.5 kb block region of *XDH* that contain the SNPs defined for the respective genes.

7. HAPLOTYPE ASSOCIATION ANALYSIS

Haplotype frequencies were estimated by Expectation Maximization (EM) algorithm; the inferred haplotypes with frequencies greater than 0.01 were compared between responders

and intolerant patients to test associations statistically with Haplo.score package as outlined by Schaid et al.[77] The computed global score and haplotype-specific p-values were calculated and adjusted sex as a clinical covariate under the additive model. The p-values were corrected for multiple testing by 10000-time permutation tests. The relationships between haplotype and clinical outcome were examined using a generalized linear model regression of trait-on-haplotype effect (haplo.glm function), which estimates regression coefficients corresponding to each haplotype.

RESULTS

1. PATIENTS' RESPONSIVENESS TO AZA

Of the 71 patients treated with azathioprine, 30 were classified as responsive to AZA, 16 as unresponsive,[73] and 25 as intolerant[72] with side effects including gastric intolerance (9/25 patients), hepatotoxicity (11/25 patients), leucopenia (1/25 patients), rash (1/25 patients), myelosuppression (1/25 patients) and others (2/25 patients). A summary of patient data is shown in **Table 3**. Steroids were given in association with AZA to 54 patients (76%); in the remaining seventeen patients (10 intolerant, 3 unresponsive and 4 responsive to AZA), AZA was administered alone because steroids were contraindicated or not tolerated.

Table 3. Patient data and responses to AZA

Response to AZA	Intolerant (n=25)	Unresponsive (n=16)	Responsive (n=30)
Sex	19/6	14/2	19/11
Age at the first AZA administration (years)	43.7±15.1 (19-67)	46.7±18 (24-81)	50±20.9 (15-85)

2. NEW INTRONIC POLYMORPHISM G420-4A IN *TPMT* GENE

Because our SNP identification strategy uses primers that, in addition to the relevant exonic sequence, allow the amplification of about 50 bases of intronic sequences flanking the exon on each side, in *TPMT* gene we found a previously undescribed nucleotidic change (G420-4A) located 4 bases upstream of exon 7 in one intolerant patient with a heterozygous genotype (**Figure 11a**). We submitted this SNP to NCBI SNPs database[78] and it has been assigned the RefSNP ID: rs56019966. We have sequenced the relevant DNA region from 100 randomly selected healthy individuals and detected this new SNP in only one individual (**Figure 11b**). Because of its position at the beginning of exon 7, the newly identified intronic polymorphism could affect the splicing process; however, prediction of the splice sites for exon 7 indicates that a mutation from G to A would not change the acceptor site or result in a new acceptor site for splicing of exon 6 to exon 7 (**Figure 11c**).

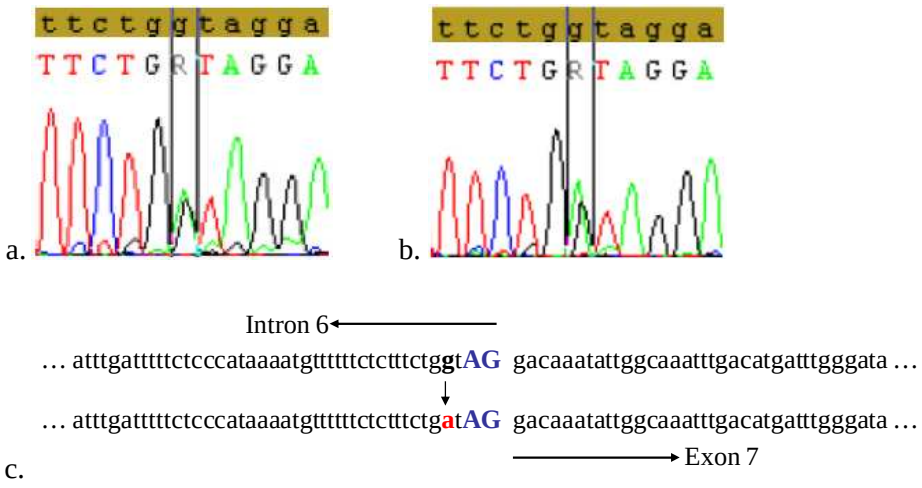


Figure 11. Identification of a new *TPMT* polymorphism (G420-4A). Chromatograms of relevant DNA sequences in (a) patient and (b) healthy control; vertical lines highlight the guanine-to-adenine mutation, present in heterozygous form (R). (c) G420-4A does not affect the acceptor site for splicing of exon 6 to exon 7. Part of the 3' end of intron 6 and at the 5' end of exon 7 are shown for the wild-type (top) and the SNP-containing (bottom) sequences. The acceptor splice site is shown in blue, the mutated nucleotide in red.

3. DISTRIBUTION OF THE *TPMT* ALLELES ACCORDING TO AZA RESPONSIVENESS

Analysis of the distribution of *TPMT* alleles into the three different groups of patients segregated according to their responsiveness, unresponsiveness, or intolerance to AZA, confirmed the presence of several known intronic and exonic *TPMT* polymorphisms that, singly, were not correlated with particular types of responses to AZA (**Table 4**), as also found in other studies. In contrast to what is generally accepted, we observed heterozygous *TPMT**3A not associated with intolerance in two patients, one unresponsive and one responsive to AZA (**Table 4**). As shown in **Table 5**, the genotype distributions of the eight *TPMT* SNPs detected in our patients are all in Hardy-Weinberg equilibrium (HWE p-value >0.05). The new SNP, G420-4A, had a minor allele frequency (MAF) <1% of the total population (**Table 5**) and was therefore excluded from further analysis.

Table 4. Distribution of *TPMT* alleles according to AZA responsiveness

Allele	AZA response	Wild-type frequency	Frequency as Heterozygous	Frequency as Homozygous
A141-101T	I	30 (60%)	14 (28%)	6 (12%)
	U	16 (50%)	6 (18.8%)	10 (31.2%)
	R	27 (45%)	15 (25%)	18 (30%)
T366+58C	I	30 (60%)	14 (28%)	6 (12%)
	U	15 (46.9%)	7 (21.9%)	10 (31.2%)
	R	26 (43.3%)	14 (23.3%)	20 (33.4%)
G420-4A[a]	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
C474T[b]	I	41 (82%)	7 (14%)	2 (4%)
	U	24 (75%)	4 (12.5%)	4 (12.5%)
	R	46 (76.7%)	12 (20%)	2 (3.3%)
C339T[b]	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
*3A (*3B+*3C)	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
*3A + T140+114A	I	46 (92%)	4 (8%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)

I: intolerant (n = 25); U: unresponsive (n = 16); R: responsive (n = 30); [a] Newly identified; [b] Synonymous polymorphism

Table 5. *TPMT* SNPs identified in our cohort of patients

SNPs	Alleles	MAF	p-val HWE	Gene Mutation	Amino acid substitution
T140+114A	T/A	0.028	1	Intronic	-
A141-101T	A/T	0.337	1	Intronic	-
C339T	C/T	0.014	1	Synonymous	None
T366+58C	C/T	0.05	1	Intronic	-
G420-4A	G/A	0.007	1	Intronic	-
G460A	G/A	0.042	1	Non Synonymous	A/T
C474T	C/T	0.218	0.857	Synonymous	None
A719G	A/G	0.042	1	Non synonymous	Y/C

Red: new intronic polymorphism

4. HAPLOTYPE ASSOCIATION ANALYSIS

Using Solid Spine method, we defined a linkage disequilibrium (LD) block spanning 18 kb for the seven relevant SNPs in the 71 patients (**Figure 12**). Strong evidence for pairwise LD was observed between T140+114A and each SNP of the *3A haplotype (G460A, A719G) with D' = 1 and $r^2 = 0.65$ (**Figure 12**).

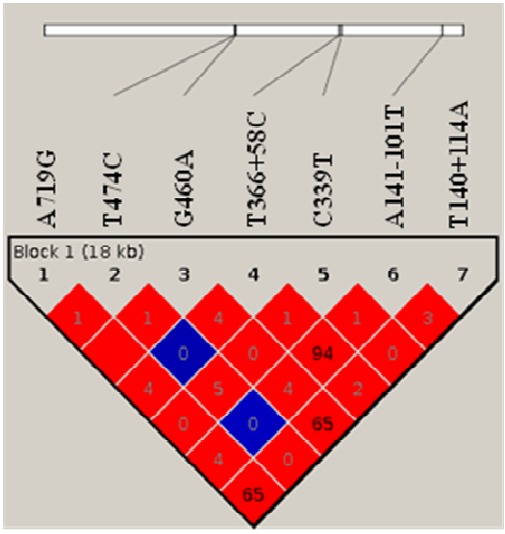


Figure 12. Haplotype block diagram of the 18 kb *TPMT* region analysed in this study. The LD plot, as determined by Solid Spine method, defines the 18 kb *TPMT* in one haplotype block. Pairwise LD among the SNPs was examined using standard coefficient D' (D' = 0: blue; D' = 1: red) and LD coefficient expressed as r^2 (numbers in rhombi).

Haplotype estimation yielded eight major haplotypes, of which five had a frequency cut-off >1% (**Table 6**). The test for overall distribution of haplotypes showed statistical differences in the distribution of haplotype frequencies between intolerant and responsive patients. The computed global score and haplotype-specific p-values were calculated and adjusted with clinical covariate (sex), under the additive model (Global simulated p-value = 0.0238). Unconditional regression analysis revealed that Hap5 [ATCCACG; Psim = 0.0372, Odds Ratio (OR) = 1.67, Confidence Interval (CI) = 1.06-2.62, **Table 7**] was significantly associated with intolerance, compared with the most frequent haplotype Hap 1 (TACTGTA; **Table 6 and Table 7**); the simulated p-value (Psim) was still significant after permutation (10000 times).

Table 6. Haplotype evaluation through Expectation Maximixation algorithm.

Hap[a]	**T140+ 114A**	**A141- 101T**	**C339T**	**T366+ 58C**	**G460A**	**C474T**	**A719G**	**R+I Freq**[b]
Hap1	T	A	C	T	G	T	A	0.430
Hap2	T	T	C	C	G	T	A	0.306
Hap3	T	A	C	T	G	C	A	0.079
Hap4	T	T	C	C	G	C	A	0.11
Hap5	**A**	**T**	**C**	**C**	**A**	**C**	**G**	**0.036**
Hap6	T	A	C	C	G	C	A	0.009
Hap7	T	T	T	C	G	C	A	0.009
Hap8	T	T	C	C	A	T	G	0.008

[a]Haplotype identified in our patients; [b]R+I Freq, haplotype frequency in intolerant and responder patient population

Of the five major haplotypes with a relevant frequency cut-off (>1%; Hap1-Hap5; **Table 7**), two, Hap2 and Hap4, had a negative Hap-Score (Hap2 = -1.5202, Hap4 = -1.69), indicating that intolerance was unlikely in the presence of these haplotypes. In contrast, the frequency of the Hap2 and Hap4 haplotypes was high in responders, albeit not significant.

Table 7. Summary of the statistics association study of relevant TPMT haplotypes

Hap[a]	R+I Freq[b]	R Freq[c]	I Freq[d]	Hap-Score[e]	p-val[f]	Psim[g]	OR[h]	CI[i]
Hap1	0.430	0.404	0.420	+1.096	0.273	0.288	Base	Base
Hap2	0.306	0.346	0.300	-1.520	0.128	0.139	0.938	0.784-1.122
Hap3	0.079	0.030	0.180	+1.308	0.191	0.220	1.430	1.071-1.905
Hap4	0.112	0.170	9.394E-07	-1.689	0.091	0.090	0.619	0.451-0.851
Hap5	**0.036**	**-**	**0.08**	**+2.254**	**0.024**	**0.037**	**1.671**	**1.065-2.622**

[a] Haplotypes; [b] R+I Freq, haplotype frequency in responder and intolerant patient population; [c] R Freq, haplotype frequency in responder patients; [d] I Freq, haplotype frequency in intolerant patients; [e]Hap-Score obtained using the haplo.score function; [f]p-val, p-value obtained using the haplo.score function; [g] Psim, p-value of 10000-time permutation test; [h] OR, Odds Ratio; [i] CI, Confidence Interval.

5. NEW INTRONIC POLYMORPHISM IN *XDH* GENE

To identify SNPs in the *XDH* gene, we have used the same strategy as that used to identify *TPMT* SNPs, with primers that allow the amplification of about 50 bases of intronic sequences flanking the exon on each side. We found one previously undescribed nucleotidic change located in intron 8. The new intronic nucleotidic change (T652-21A) is located 21 bases upstream of exon 9 in one intolerant patient with a heterozygous genotype (**Figure 13**).

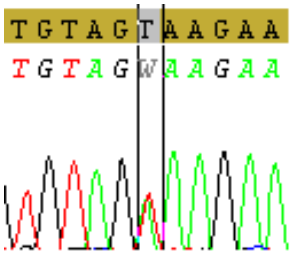


Figure 13. Identification of a new intronic *XDH* polymorphism. (a) Chromatogram of the relevant DNA sequence in the intolerant patient; vertical lines highlight the thymine-to-adenine mutation, present in heterozygous form (W).

6. NEW EXONIC MUTATIONS IN *XDH* GENE

By genotyping all the 36 exons of *XDH* gene, we identified three new mutations in exons 11, 13, and 26, and one polymorphism in exon 16.

In exon 11, we found a non-synonymous nucleotidic change c.G1004A, p.Arg335His present in heterozygous form in one intolerant patient. This mutation is located in the FAD-binding domain (**Figure 14a**). The occurrence of this mutation was sought, but not found, in 100 healthy controls. It is interesting to note that Arg at position 335 is conserved in higher species (**Figure 14b**), suggesting that it might play an important role in the protein; hence, a change from Arg to the cyclic His is likely to have a significant effect. Indeed, the FoldX algorithm revealed a positive difference in free energy of the mutated protein (2.8 kcal/mol), indicating that c.G1004A, p.Arg335His destabilizes the structure of XO (**Figure 14c** and **Table 9**).

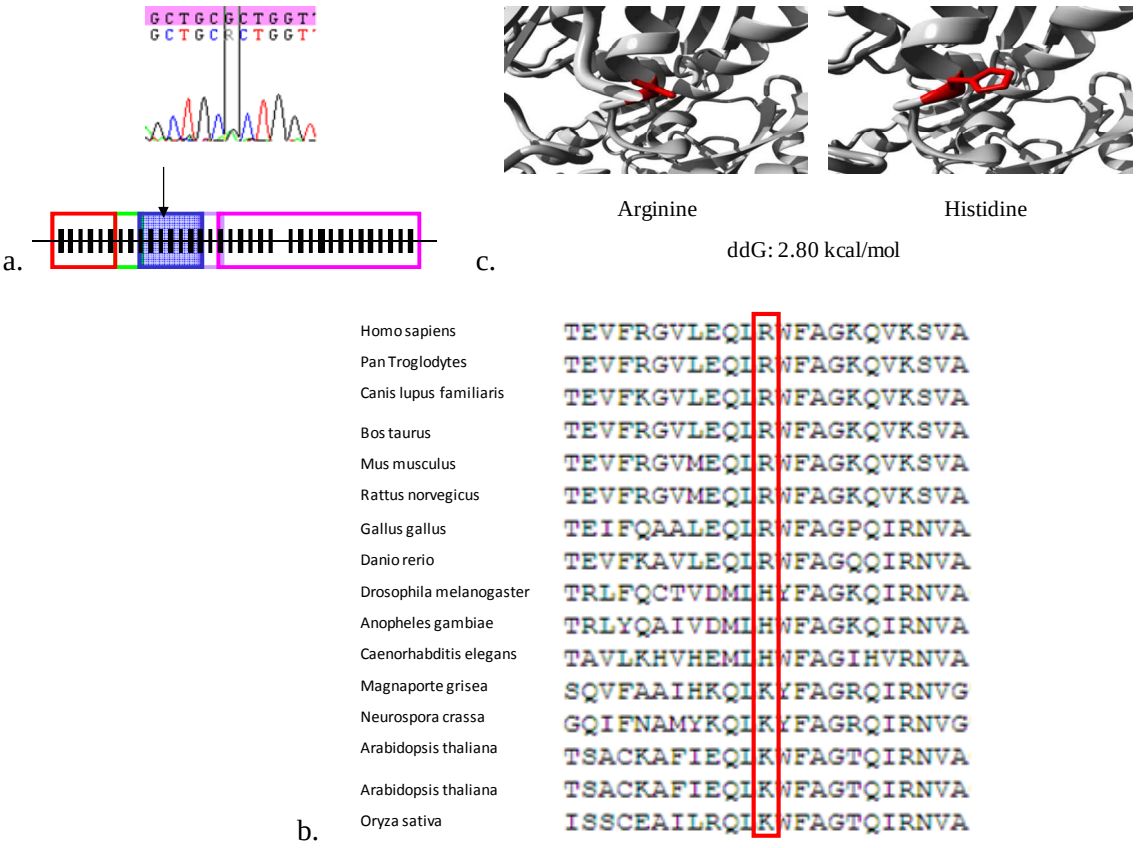


Figure 14. Characteristics of the c.G1004A,p.Arg335His mutation. a. Chromatogram of the relevant DNA region encompassing the c.G1004A sequence; vertical lines highlight the guanine-to-adenine variant, present in heterozygous form (R); the mutation (arrow) is located in the FAD binding domain (blue block); b. Arginine (R) is conserved in *XHD* orthologs from many species; c. 3D representation of wild-type and mutated XO structure focused on the area of the mutation; the ddG value is indicated.

In exon 13, we found a synonymous nucleotidic change c.G1194A, p.Leu398Leu, located in the sequence coding for the FAD-binding domain (**Figure 15**), in one intolerant patient; hence, the mutation has no effect on the protein product (**Table 9**). The occurrence of this mutation was sought, but not found, in 100 healthy controls.

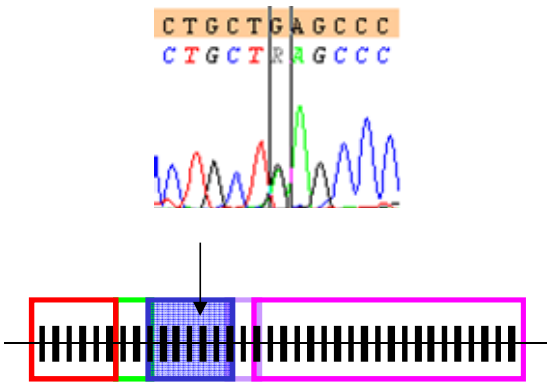


Figure 15. Characteristics of the c.G1194A, p.Leu398Leu mutation. Chromatogram of the relevant DNA region encompassing the c.G1194A sequence; vertical lines highlight the guanine-to-adenine mutation, present in heterozygous form (R); the mutation (arrow) is located in the FAD-binding domain (blue block).

In exon 16, we found the non-synonymous nucleotidic change c.C1167T, p.Ala556Val present in heterozygous form in one intolerant, one unresponsive, and two responsive patients. This polymorphism is located in the connection segment between the FAD-binding domain and the molybdopterin domain (**Figure 16a**) and Ala at position 556 is conserved in mammalian species (**Figure 16b**). It is interesting to note that the occurrence of this polymorphism was found in 3/100 healthy controls, suggesting that it might not play an important role in the protein. Indeed, the FoldX algorithm showed a weakly negative difference in free energy of the modified protein (-0.48 kcal/mol), indicated that c.C1167T, p.Ala556Val, does not destabilize the structure of XO (**Figure 16c** and **Table 9**).

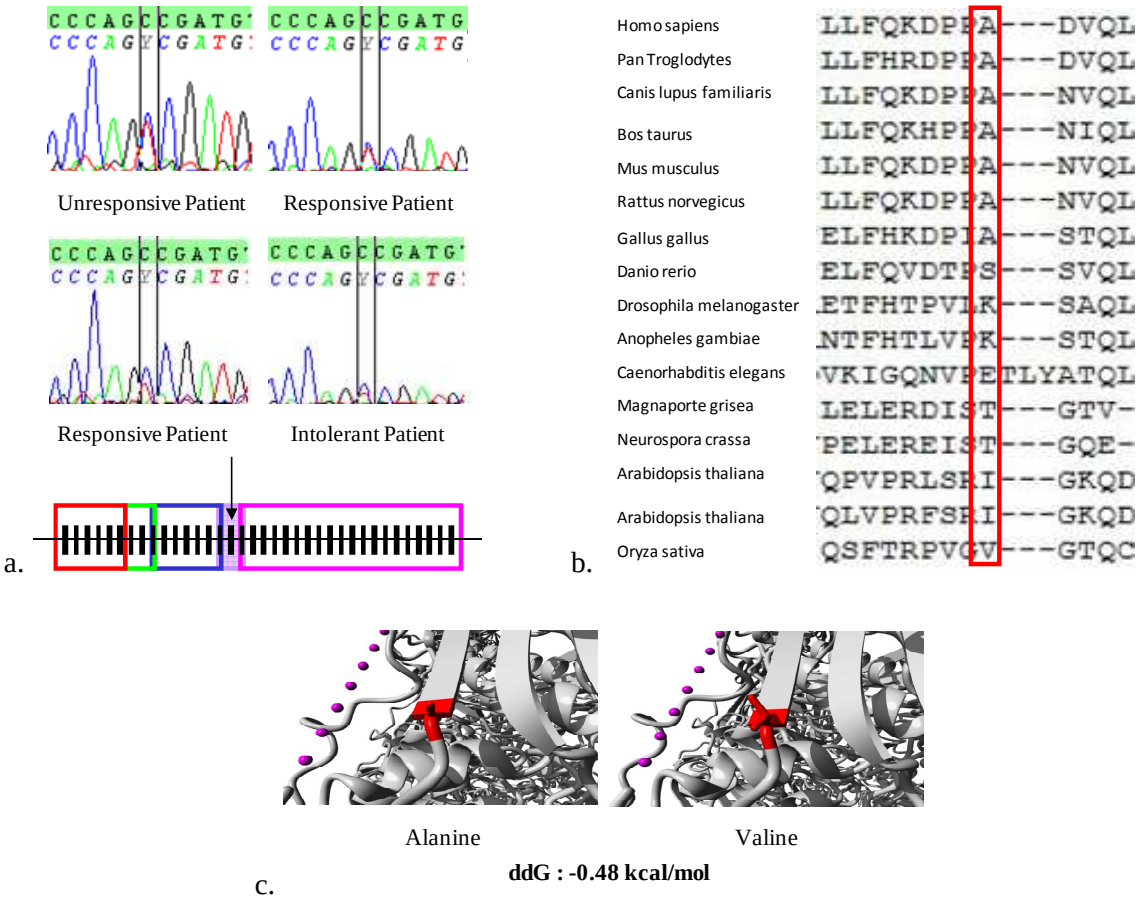


Figure 16. Characteristics of c.C1167T, p.Ala556Val mutation. a. Chromatogram of relevant DNA region encompassing c.C1167T sequence; vertical lines highlight the cytosine-to-thymine polymorphism, present in heterozygous form (Y). The polymorphism is located in exon 16 (arrow) encoding part of the sequence of the connecting segment between the FAD-binding domain and the molybdopterin domain (purple block); b. Alanine (A) is conserved in *XDH* orthologs from many species; c. 3D representation of wild-type and mutated XO structure focused on the area of the mutation; the ddG value is indicated.

In exon 26, we found a non-synonymous nucleotiditic change c.C2891T, p.Thr964Ile present in heterozygous form, in one responsive patient. This mutation is in the molybdopterin domain (**Figure 17a**). The occurrence of this mutation was sought, but not found, in 100 healthy controls. It is interesting to note that Thr at position 964 is conserved in higher species (**Figure 17b**), suggesting that it might play an important role in the protein, hence, a change from Thr to the Ile is likely to have a significant effect. Indeed, the FoldX algorithm revealed a positive difference in free energy of the mutated protein (2.06 kcal/mol), indicating that c.C2891T, p.Thr964Ile, destabilizes the structure of the XO (**Figure 17c** and **Table 9**).

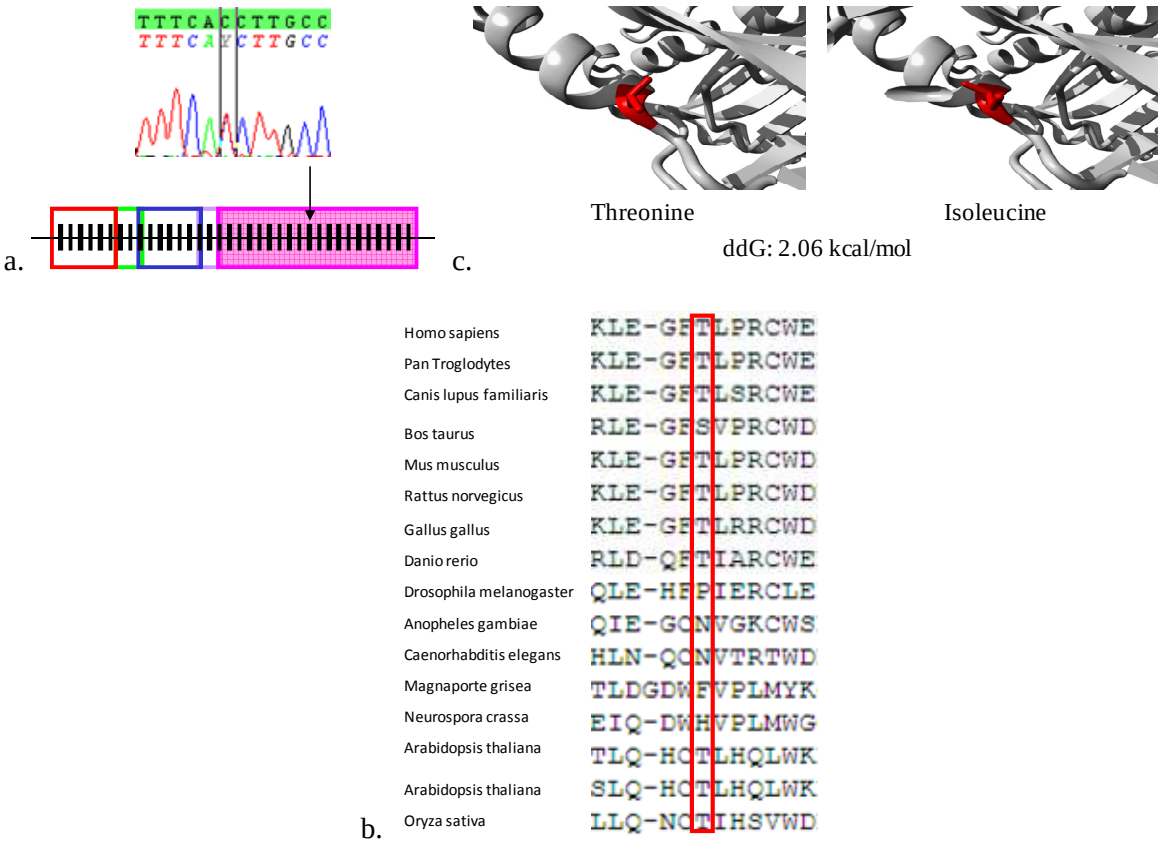


Figure 17. Characteristics of c.C2891T, p.Thr964Ile. a. Chromatogram of the relevant DNA region encompassing c. C2891T; vertical lines highlight the cytosine-to-thymine variant, present in heterozygous form (Y) the mutation (arrow) is located in the molybdopterin domain (pink block); b. threonine (T) is conserved in *XHD* orthologs from many species; c. 3D representation of wild-type and mutated XO structure focused on the area of the mutation; the ddG value is indicated.

7. DISTRIBUTION OF THE *XDH* ALLELES ACCORDING TO AZA RESPONSIVENESS

Studies of the distribution of *XDH* allele into the three different groups of patients divided according to their responsiveness, unresponsiveness, or intolerance to AZA, verified the presence of some known intronic and exonic *XDH* polymorphisms that, singly, were not associated with specific types of responses to AZA (**Table 8**). We observed heterozygosity for A2107G (**Table 8**), a polymorphism shown by Kudo et al. to determines hyper-activation of XO,[47] albeit with no statistically significant correlation, and thereby increased metabolism of 6-MP to thiouric acid and, hence, decreased production of 6TGNs from AZA. Accordingly, the A2107G polymorphism should lead to unresponsiveness to AZA. In our cohort, however, this polymorphism was identified not only in 4 unresponsive patients, but also in one responsive and two intolerant patients. All 7 patients, regardless of

their status in response to AZA, show the polymorphism A2107G in association with another non-synonymous polymorphism, A1936G (**Table 8**).

Table 8. Distribution of *XDH* alleles according to AZA responsiveness

Allele	AZA response	Wild-type frequency	Frequency as Heterozygous	Frequency as Homozygous
C100+27T	I	47 (94%)	3 (6%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
A307-33G	I	45 (90%)	5 (10%)	0 (0%)
	U	26 (81.3%)	6 (18.75%)	0 (0%)
	R	51 (85%)	7 (11.7%)	2 (3.3%)
G514A	I	48 (96%)	2 (4%)	0 (0%)
	U	30 (93.8%)	2 (6.2%)	0(0%)
	R	53 (88.3%)	7 (11.7%)	0 (0%)
G627A[b]	I	48 (96%)	2 (4%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	54 (90%)	6 (10%)	0 (0%)
T652-21A[a]	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
G822A[b]	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
C837T[b]	I	48 (96%)	2 (4%)	0 (0%)
	U	30 (93.8%)	2 (6.2%)	0(0%)
	R	54 (90%)	6 (10%)	0 (0%)
C859G	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
G1004A[a]	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
G1194A[a]	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
C1274G	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
G1329A[b]	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	58 (96.7%)	2 (3.3%)	0 (0%)
C1509G[b]	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
C1667T[a]	I	49 (98%)	1 (2%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	58 (96.7%)	2 (3.3%)	0 (0%)
G1857-4A	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)

A1936G	I	48 (96%)	2 (4%)	0 (0%)
	U	28 (87.5%)	4 (12.5%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
A2107G	I	48 (96%)	2 (4%)	0 (0%)
	U	28 (87.5%)	4 (12.5%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
G2197+42C	I	37 (74%)	9 (18%)	4 (8%)
	U	28 (87.5%)	4 (12.5%)	0 (0%)
	R	41 (68.3%)	11 (18.3%)	8 (13.4%)
G2197+68A	I	31 (62%)	13 (26%)	6 (12%)
	U	19 (59.4%)	9 (28.1%)	4 (12.5%)
	R	46 (76.7%)	10 (16.7%)	4 (6.6%)
C2211T[b]	I	38 (76%)	8 (16%)	4 (8%)
	U	29 (90.6%)	3 (9.4%)	0 (0%)
	R	46 (76.7%)	8 (13.3%)	6 (10%)
A2544+35T	I	10 (20%)	6 (12%)	34 (68%)
	U	8 (25%)	6 (18.7%)	18 (56.3%)
	R	23 (38.3%)	13 (21.7%)	24 (40%)
G2824-35A	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	58 (96.7%)	2 (3.3%)	0 (0%)
C2891T[a]	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
T3030C[b]	I	12 (24%)	8 (16%)	30 (60%)
	U	10 (31.2%)	8 (25%)	14 (43.8%)
	R	17 (28.3%)	13 (21.7%)	30 (50%)
C3052-12G	I	13 (26%)	7 (14%)	30 (60%)
	U	10 (31.2%)	8 (25%)	14 (43.8%)
	R	16 (26.7%)	12 (20%)	32 (53.3%)
A3352-21C	I	49 (98%)	1 (2%)	0 (0%)
	U	30 (93.8%)	2 (6.2%)	0 (0%)
	R	54 (90%)	2 (3.3%)	4 (6.7%)
T3488C	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
T3536C	I	50 (100%)	0 (0%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	60 (100%)	0 (0%)	0 (0%)
G3717A[b]	I	49 (98%)	1 (2%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
C3774+32G	I	49 (98%)	1 (2%)	0 (0%)
	U	31 (96.9%)	1 (3.1%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
C3886T	I	50 (100%)	0 (0%)	0 (0%)
	U	32 (100%)	0 (0%)	0 (0%)
	R	59 (98.3%)	1 (1.7%)	0 (0%)
A4002+518G	I	20 (40%)	10 (20%)	20 (40%)
	U	22 (68.8%)	8 (25%)	2 (6.2%)
	R	36 (60%)	14 (23.3%)	10 (16.7%)
A4002+1477G	I	38 (76%)	10 (20%)	2 (4%)
	U	20 (62.5%)	8 (25%)	4 (12.5%)
	R	40 (66.7%)	14 (23.3%)	6 (10%)

C4002+1516T	I	20 (40%)	12 (24%)	18 (36%)
	U	22 (68.8%)	8 (25%)	2 (6.2%)
	R	39 (65%)	11 (18.3%)	10 (16.7%)
G4002+1555T	I	40 (80%)	8 (16%)	2 (4%)
	U	24 (75%)	6 (18.7%)	2 (6.3%)
	R	43 (71.7%)	9 (15%)	8 (13.3%)

I: intolerant (n = 25); U: unresponsive (n = 16); R: responsive (n = 30); [a] Newly identified; [b] Synonymous polymorphism

As shown in **Table 9**, the genotype distributions of the thirty-five *XDH* nucleotidic changes detected in our patients, including the four new mutations and polymorphism we have demonstrated, are all in Hardy-Weinberg equilibrium (HWE p-value >0.05), with exception of the intronic A3352-30C, which was excluded from subsequent analysis. Of these 35 changes, the following 12 SNPs and mutations, T652-21A, G822A, C859G, G1004A, G1194A, C1274G, C1509G, G1857-4A, C2891T, T3488C, T3536C, and C3886T, were also excluded from further analysis as they had a minor allele frequency (MAF) <1% of the total population (**Table 9**).

Table 9. *XDH* nucleotidic changes identified in our cohort of patients

SNPs	Allele	MAF	p-Val HWE	Gene mutation	Amino acid substitution	Protein destabilization
C100+27T	C/T	0.035	1	Intronic	-	-
A307-33G	A/G	0.141	1	Intronic	-	-
G514A	G/A	0.077	1	Non Synonymous	G/R	-
G627A	G/A	0.063	1	Synonymous	None	-
T652-21A	T/A	0.007	1	Intronic	-	-
G822A	G/A	0.007	1	Synonymous	None	-
C837T	C/T	0.07	1	Synonymous	None	-
C859G	C/G	0.007	1	Non Synonymous	L/V	-
G1004A	G/A	0.007	1	Non Synonymous	R/H	Yes
G1194A	G/A	0.007	1	Synonymous	None	No
C1274G	C/G	0.007	1	Non Synonymous	S/C	-
G1329A	G/A	0.014	1	Synonymous	None	-
C1509G	C/G	0.007	1	Synonymous	None	-
C1667T	C/T	0.028	1	Non Synonymous	A/V	Yes
G1857-4A	G/A	0.007	1	Intronic	-	-
A1936G	A/G	0.049	1	Non Synonymous	I/V	-
A2107G	A/G	0.049	1	Non Synonymous	I/V	-
G2197+42C	G/C	0.254	0.51	Intronic	-	-
G2197+68A	G/A	0.324	1	Intronic	-	-
C2211T	C/T	0.204	0.24	Synonymous	None	-
A2544+35T	T/A	0.289	0.33	Intronic	-	-
G2824-35A	G/A	0.014	1	Intronic	-	-

C2891T	C/T	0.007	1	Non Synonymous	T/I
T3030C	C/T	0.275	1	Synonymous	None
C3052-12G	G/C	0.275	0.87	Intronic	-
A3352-30C	A/C	0.063	0.04	Intronic	-
T3488C	T/C	0.007	1	Non Synonymous	V/A
T3536C	T/C	0.007	1	Non Synonymous	I/T
G3717A	G/A	0.014	1	Synonymous	None
C3774+32G	C/G	0.021	1	Intronic	-
C3886T	C/T	0.007	1	Non Synonymous	R/W
A4002+518G	A/G	0.451	0.56	3'UTR	-
A4002+1477G	A/G	0.31	0.92	3'UTR	-
C4002+1516T	C/T	0.43	0.46	3'UTR	-
G4002+1555T	G/T	0.246	0.40	3'UTR	-

In red: newly detected mutations and polymorphisms

8. HAPLOTYPE ASSOCIATION ANALYSIS

Using the Solid Spine method, we defined a four LD blocks spanning 71.5 kb for the 21 relevant SNPs in the 71 patients (**Figure 18**).

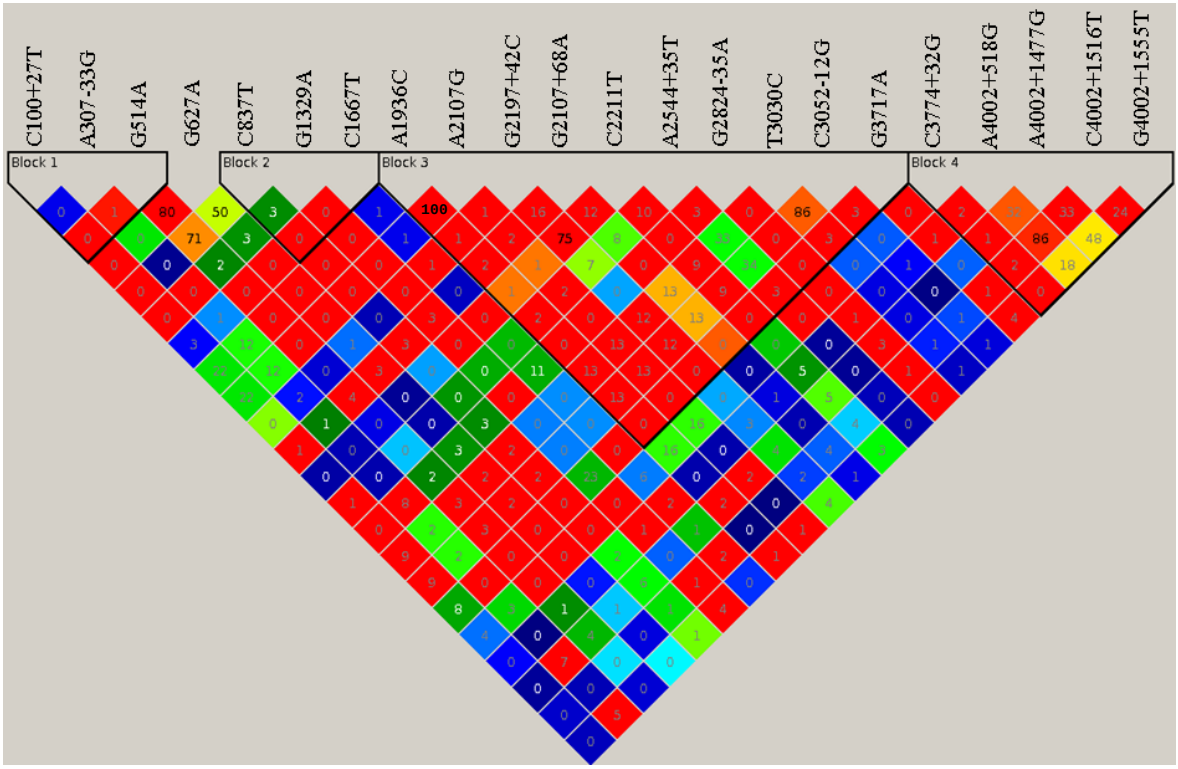


Figure 18. Haplotype block diagram of the 71.5 kb *XDH* gene region analyzed in this study. The LD plot, as determined by Solid Spine method, defines the 71.5 kb *XDH* in four haplotype blocks. Pairwise LD among the SNPs was examined using standard coefficient D' (D' = 0: blue; D' = 1: red; 0> D' <1: other colours) and LD coefficient expressed as r^2 (numbers in rhombi).

The Spine algorithm defined four LD blocks within the 71.5 kb *XDH* gene region analyzed in this study. Haplotype estimation showed that, of all four LD blocks, only block 4 contains a haplotype significantly associated with intolerance. Thus, haplotype evaluation yielded seven major haplotypes within block 4, as per a frequency cut-off >1% (**Table 10**); of these, unconditional regression analysis revealed that Hap6 (CGATG; PSim = 0.017, **Table 11**) was significantly associated with intolerance, compared to the second most frequent haplotype Hap4 (CAGCT; **Table 10** and **Table 11**); Psim was still significant after permutation (10000 times). The computed global score and haplotype-specific p-values were calculated and adjusted with clinical covariate (sex), under the additive model (Global simulated p-value = 0.241).

Of the seven major haplotypes estimated (Hap1-Hap7; **Table 10**), two, Hap3 and Hap5, had a negative Hap-Score (cut-off > -1; Hap3 = -1.257, Hap5 = -1.479), indicating that intolerance was unlikely in the presence of these haplotypes. In contrast, the frequencies of the Hap3 and Hap5 haplotypes were high in responders, albeit not significant.

Table 10. Haplotype evaluation through Expectation Maximixation algorithm.

Hap[a]	C3774+ 32G	A4002+ 518G	A4002+ 1477G	C4002+ 1516T	G4002+ 1555T	R+I Freq[b]
Hap1	C	A	A	C	G	0.198
Hap2	C	A	A	C	T	0.021
Hap3	C	A	G	C	G	0.074
Hap4	C	A	G	C	T	0.206
Hap5	C	G	A	C	G	0.018
Hap6	**C**	**G**	**A**	**T**	**G**	**0.445**
Hap7	C	G	G	C	T	0.010

[a]Haplotype identified in our patients; [b]Frequency of haplotype in our cohort

Table 11. Summary of statistics association study of XDH haplotypes

Hap[a]	R+I Freq[b]	R Freq[c]	I Freq[d]	Hap-Score[e]	p-val[f]	Psim[g]	OR[h]	CI[i]
Hap1	0.198	0.230	0.160	-0.909	0.363	0.377	1.087	0.815-1.450
Hap2	0.021	0.037	-	-0.972	0.331	0.609	0.871	0.520-1.458
Hap3	0.074	0.103	0.040	-1.257	0.209	0.190	0.844	0.566-1.258
Hap4	0.206	0.230	0.180	-0.636	0.525	0.603	Base	Base
Hap5	0.018	0.033	-	-1.479	0.139	0.208	0.702	0.340-1.445
Hap6	**0.445**	**0.333**	**0.580**	**+2.391**	**0.017**	**0.017**	**1.229**	**0.968-1.561**
Hap7	0.010	-	0.020	+0.978	0.328	0.658	1.636	0.600-4.460

[a] Haplotypes; [b] R+I Freq, haplotype frequency in responder and intolerant patient population; [c] R Freq, haplotype frequency in responder patients; [d] I Freq, haplotype frequency in intolerant patients; [e] Hap-Score obtained using the haplo.score function; [f] p-val, p-value obtained using the haplo.score function; [g] Psim, p-value of 10000-time permutation test; [h] OR, Odds Ratio; [i] CI, Confidence Interval.

DISCUSSION

Azathioprine is one of the most widely used immunosuppressant drugs for the treatment of many different diseases such as inflammatory bowel disease, Crohn's disease, ulcerative colitis,[3] rheumatologic diseases,[2] rheumatoid arthritis,[79] myasthenia gravis[80,81] and multiple sclerosis.[82] In autoimmune diseases, AZA is often administered as a steroid-sparing drug.[81,83-86] However, responsiveness to the drug differs in different patients, and the possibility to predict unresponsiveness and/or intolerance is of outmost importance for safe and effective treatment. There is increasing evidence from pharmacogenetics studies that hereditary changes in genes coding for TPMT and XO can affect the response to AZA,[87] by altering the functionality of these enzymes or determining their rapid degradation.[88] Towards further delineation of genetic *TPMT* and *XDH* polymorphisms that might be specifically associated with the different types of response to AZA and could thereby be used as part of pharmacogenetic profiles to identify intolerant and/or unresponsive patients, we have genotyped 71 Italian patients treated with AZA.

Through these studies, we uncovered new intronic polymorphisms, one in *TPMT*, G420-4A, and one in *XDH*, T652-21A. Each polymorphism was carried by one different AZA-intolerant patient, with *TPMT* G420-4A also detected in one of 100 healthy controls investigated. As these nucleotidic changes are located four bases upstream of the RNA splicing site for exon 7 of *TPMT* gene and 21 bases upstream of the RNA splicing site for exon 9 of *XDH* gene, we hypothesize that they could affect mRNA splicing,[89,90] thus coding for truncated proteins with defective enzymatic activity. However, G420-4A splice site prediction indicated that the polymorphism does not affect the relevant acceptor site, nor does it create a new splicing site. Nevertheless, we cannot exclude that *TPMT* G420-4A could interfere with the protein of the spliceosome that binds upstream of exon 7 in the region of the acceptor site.[91-93] Similarly, *XDH* T652-21A could affect the association of the small nuclear RNA–protein complex U2 with the branch point region of the intron, located 20 - 50 bases upstream of the splice acceptor site;[94,95] both polymorphisms therefore have the potential to affect mRNA splicing and the protein product. Since the variants (*TPMT* G420-4A and *XDH* T652-21A) were each found only in one AZA-treated patient, further studies in larger cohorts of varying ethnicities are necessary to evaluate their frequency in the general population. Such studies would also permit to better

understand if intolerance to AZA in the patients was indeed related to alteration of the TPMT or XO enzymes imparted by these nucleotidic changes.

Through genotyping of the *TMPT* gene, we detected a novel haplotype defined by seven SNPs, which include both SNPs of *TPMT**3A and which we have termed *TPMT**3E. *TPMT**3E was significantly associated with intolerance in our population (**Table 7**, Psim = 0.037), and none of the 16 unresponsive or 30 responsive patients carried this novel haplotype. The T140+114A polymorphism was found in our cohort only in association with the *TPMT**3A mutations, as part of *TPMT**3E (4/25 intolerant patients); none of the patients carried this SNP on its own. At present, it is not possible to ascertain if the presence of the T140+114A polymorphism amplifies the intolerant response to the drug that is associated with mutations encompassed in *TPMT**3A. Studies on larger cohorts will be needed to determine whether this SNP is always found as part of the *TPMT**3E haplotype or also occurs alone, and if so, whether it could by itself be associated with intolerance; in this context, the study of Otterness et al., who first reported this SNP, is not clear.[38] Larger, diverse, cohorts are also necessary to investigate ethnic variations in frequency of *TPMT**3E or the single T140+114A polymorphism. Indeed, the frequency of *TPMT* alleles in different ethnic groups varies greatly; for example, *TPMT**3A, which is the most prominent allele in Caucasians, was not found in Ghanaian and Korean individuals,[96-98] and the frequency of *TPMT**3C, the most prominent allele in Ghanaian and Korean individuals, is very low in Caucasians.[99] We could only find reports for T140+114A frequencies for two different populations in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=3931660): 6% for a Nigerian cohort of 25 individuals and 9.4% for a cohort of 90 individuals whose ethnic origin is not reported. It is therefore difficult to discuss if the allele frequency for T140+114A in our AZA-treated Italian cohort (3.2%) is commensurate with that of other populations. Strikingly, whereas the *TPMT**3A haplotype was seen also in the context of unresponsiveness or responsiveness in our cohort, none of the unresponsive or responsive patients carried the *TPMT**3E haplotype. This could suggest a very strong link of *TPMT**3E with AZA intolerance. The *TPMT**3A haplotype has been linked to intolerance to AZA[87] and appears to be the most frequent inactivating mutant allele in Caucasians (3.2-5.7%),[96,100,101] including Italian subjects (3.9%).[102] In our cohort, the percentage of patients heterozygous for *TPMT**3A mutations was over twice as high (6/71 patients; 8.5%), a difference that could be related to the small cohort studied. *TPMT**3A has not so far been reported in association with responsiveness or unresponsiveness. In our study, however,

the *TPMT**3A haplotype was detected in one unresponsive patient and in one apparently responsive patient by DNA sequencing (**Table 4**). This responsive *TPMT**3A patient was given AZA as a steroid-sparing agent. Whether or not steroid-sparing in this *TPMT**3A patient was directly correlated to AZA treatment is not absolutely established. Nevertheless, after eleven years of AZA treatment, the patient's daily steroid dose had been reduced to over 80% from the initial dose. Neither this patient, nor the unresponsive one, experienced side effects as might be expected in individuals carrying the *TPMT**3A allele.[32] The unexpected occurrence of *TPMT**3A in association with unresponsiveness or responsiveness in our cohort could be related to one or more additional unknown variants in *TPMT* or in other genes of AZA metabolism that would abrogate the effect of the *TPMT**3A allele and, although such an effect has never been demonstrated, could lead to hyperactivity of the enzyme in the unresponsive patient.

Similarly, the *XDH* A2107G polymorphism, which has been associated with hyper-activation of XO[47] and should therefore lead to unresponsiveness to AZA, was detected in two intolerant and one responsive patients, in addition to four unresponsive patients. LD analysis showed that this mutation is strongly associated with A1936G (r^2=100) in all these seven patients; indeed, neither mutation was found in the other 64 patients. Unfortunately, how the *XDH* A1936G polymorphism affects XO activity is not known, nor are possible associations of A1936G with other *XDH* polymorphisms. Obviously, however, as all seven patients with different responses to AZA bore both polymorphisms, it is unclear if either, both, or neither affect response to AZA; indeed, the different types of responses could also be due to other single genetic changes or combinations with *XDH* A2107G and/or A1936G polymorphisms (see below), and/or polymorphisms in other relevant genes. Unexpected effect of polymorphisms in *TPMT* or *XDH* could also be mediated at genomic level via epistasis by modifier genes.[103,104] Another possibility would involve a variation in *GST* resulting in reduced or increased enzymatic activity, and thereby only little or high conversion of AZA to 6-MP, hence unresponsiveness or intolerance, irrespective of mutations that might affect TPMT and XO downstream of GST in the pathway. Finally, unresponsiveness could be accounted for by variation(s) in *AOX1* coding for aldehyde oxidase, which contributes to the catabolism of AZA, MP, and other thiopurine metabolites. Indeed, the presence of a coding-region SNP in *AOX1* is significantly associated with lack of therapeutic response to AZA, which in this case is apparently not related to differences in TGN levels.[99]

Through genotyping of the XDH gene, we detected two novel non-synonymous mutations (c.G1004A, p.Arg335His; c.C2891T, p.Thr964Ile), one new synonymous mutation (c.G1194A, p.Leu398Leu) and one novel non-synonymous polymorphism (c.C1167T, p.Ala556Val). The first new missense mutation, c.G1004A, p.Arg335His, identified in one intolerant patient was located in the FAD-binding domain. This domain is one of the three active sites of XO protein.[58] A change of Arg for His at position 335, which is conserved in higher species, is likely to have an effect on the structural conformation of the FAD-binding domain. Indeed, while both are hydrophilic and basic amino acids, histidine contains an imidazole group, an aromatic ring that can lead to steric hindrance, thereby destabilizing the protein structure. Such possible effect of the mutation is confirmed by the positive score obtained with FoldX algorithm[105] predicting instability of the mutated XO protein. This could therefore lead to decreased XO activity, for example through increased protein degradation or through deficient binding of FAD, hence reduced efficiency of the enzyme function, and thereby intolerance to azathioprine.

The second novel non-synonymous mutation, c.C2891T, p.Thr964Ile, was found in one responsive patient (**Table 8**). This mutation was located in the molybdopterin domain, another active site of XO protein.[58] The mutation results in a change of Thr, a hydrophilic amino acid conserved in higher species, to Ile, a hydrophobic amino acid, at position 964. The replacement of a hydrophilic amino acid with a hydrophobic amino acid is likely to have a profound effect on the molybdopterin domain structure and consequently on XO activity. This was confirmed by the positive score obtained with FoldX algorithm predicting instability of the mutated XO protein. Unexpectedly, the c.C2891T, p.Thr964Ile mutation occurred in association with responsiveness, suggesting that the destabilizing effect on the protein did not in fact affect the enzymatic activity; indeed, it is likely that, in this case, the structural change imparted by the mutation does not lead to release of the molybdenum ion. Alternatively, we cannot exclude that responsiveness in the presence of a potentially destabilizing mutation could be related to one or more additional unknown variants in other genes of AZA metabolism that would abrogate the effect of this mutation.[103]

The non-synonymous polymorphism, c.C1167T, p.Ala556Val, uncovered in one intolerant, one unresponsive, two responsive patients (**Table 8**), and 3/100 healthy controls causes an amino acid change in the connection segment between the FAD-binding and molybdopterin domains, a polymorphic domain of the protein. The change of Ala to Val at position 556 is unlikely to have a significant effect on the protein, and thereby should not

alter XO activity. Indeed, Ala and Val are small, hydrophobic amino acids that can generally substitute for each other without affecting the physico-chemical characteristics of a protein. This was confirmed by FoldX algorithm calculation, which gave a weakly negative score for the mutated protein. Finally, the silent mutation, c.G1194A, p.Leu398Leu, identified in an intolerant patient should not cause functional impairment of XO. These predictions need to be validated through studies of the enzymatic function of XO in individuals with the relevant mutations.

LD analysis of *XDH* identified 4 different blocks, one of which was significantly associated with intolerance in our population (**Table 11** Psim = 0.017). It includes five SNPs, one intronic and four located in the 3' untranslated regions (3' UTRs). As the 3'UTR is important for RNA processing, one possible effect of the SNPs would be through interference in polyadenylation and thereby mRNA maturation. Processing of the pre-mRNA involves post-transcriptional cleavage of the 3′ ends that occurs at a specific site located 15 – 30 nucleotides downstream of a cleavage signal, with the sequence, AAUAAA.[106] Addition of the polyA tail resulting in a mature mRNA occurs at the cleavage site. However, in *XDH* pre-mRNA, one SNP found in the 3' UTRs is located 20 base pairs upstream of the cleavage signal AAUAAA and the other three SNPs are located around 900 - 1000 base pairs downstream of the cleavage site; accordingly, these SNPs are contained one before cleavage signal and three in the cleaved region and are therefore unlikely to have any effect on polyadenylation and thereby on mRNA maturation. Another possibility could be that the SNPs located in the 3' UTR affect the regulation of protein expression through alteration of the target sites for microRNAs, small ribonucleic acid molecules, that interact with the 3' UTRs to regulate the expression of mRNAs.[107] This could therefore lead to AZA intolerance in our patients, through in an increase in down-regulation of *XDH* mRNA and thereby XO expression. Obviously, such a hypothesis needs to be explored experimentally.

While the association of the new *TPMT**3E and *XDH* 3'UTR haplotype with intolerance to AZA must be further delineated, these studies confirm the relevance of establishing pharmacogenetic profiles for specific drugs, especially those that can result in severe side effects, to determine their suitability in individual patients. Indeed, through such analysis, we could make preliminary distinction between the seven patients who showed diverse responses to AZA, albeit bore the same *XHD* SNP (A2107G associated with A1936G) that has been associated with hyperactivity of XO, and should therefore all have shown unresponsiveness to AZA (**Table 12**). Thus, as shown in **Table 12**, while none of the

patients bore the *TPMT**3E haplotype, they could be differentiated according to presence or absence of four particular SNPs and/ or the *XDH* 3' UTR haplotype (**Table 12**). Thus, the four AZA-unresponsive patients did not have the *XDH* 3' UTR haplotype, were heterozygous for the intronic A307-33G, and wild-type for C1667T, C2211T and the intronic G2197+42C; the AZA-responsive patient, while homozygous for the *XDH* 3' UTR haplotype, was wild-type for A307-33G and C1667T, and heterozygous for G2197+42C and C2211T. The two intolerant patients differed in their profiles, with one patient homozygous for the *XDH* 3' UTR haplotype and wild-type for all four SNPs, while the other patient, who did not bear the *XDH* 3' UTR haplotype, differed from all other six patients by heterozygosity of C1667T. Obviously, further studies with larger cohorts are required to confirm and establish these profiles in the context of specific responses to AZA.

Table 12: Study of *TPMT* haplotype, *XDH* SNPs in 7 patients with polymorphism caused XO hyper-activation.

P[a]	AZA	TPMT	XDH						
	Resp	*3E	A1936G	A2107G[c]	3'UTR Hap[c]	C1667T[d]	A307 -33G	G2197 +42C	C2211T[e]
P1	U	-	Het	Het	-	wt	Het	wt	wt
P2	U	-	Het	Het	-	wt	Het	wt	wt
P3	U	-	Het	Het	-	wt	Het	wt	wt
P4	U	-	Het	Het	-	wt	Het	wt	wt
P5	R	-	Het	Het	+	wt	wt	Het	Het
P6	I	-	Het	Het	+	wt	wt	wt	wt
P7	I	-	Het	Het	-	Het	Het	Het	Het

I: intolerant; U: unresponsive; R: responsive; wt: wild-type; Het: heterozygous; +: presence of the haplotype; -: absence of haplotype; [a] P: patient; [b] Mutation causing XO hyper-activation; [c] 3'UTR Hap: 3'UTR haplotype; [d] New mutation identified in this study; [e] Synonymous mutation.

The data obtained in the studies presented in this thesis should be analyzed statistically for both *TPMT* and *XDH* genes together, to establish if profiles that incorporate data from both genes are superior to single SNP analysis in predicting types of response to AZA. In addition, and using larger cohorts, pharmacogenetic profiling for AZA should not only include studies on *TPMT* and *XDH*, but also incorporate studies on all the other genes of the AZA metabolic pathways that could play a role in the response to AZA, in particular the genes coding for GST and aldehyde oxidase, which, respectively, convert AZA into 6-MP and detoxify the drug.[99,108,109] Indeed, these, which are key enzymes in determining responsiveness, unresponsiveness or intolerance, represent future targets for our pharmacogenetics studies of the response to AZA.

REFERENCES

1. Elion GB. Symposium on immunosoppressive drugs. Biochemistry and pharmacology of purine analogues. *Fed Proc.* 1967;**26**(3):898-904.

2. Maltzman JS, Koretzky GA. Azathioprine: old drug, new actions. *J. Clin. Invest.* 2003; **111**(8):1122-1124.

3. Sahasranaman S, Howard D, Roy S. Clinical pharmacology and pharmacogenetics of thiopurines. *Eur. J. Clin. Pharmacol.* 2008;**64**(8):753-767.

4. Kothari MJ. Myasthenia gravis. *J. Am. Osteopath. Assoc.* 2004;**104**(9):377-384.

5. Mantegazza R, Antozzi C, Peluchetti D, et al. Azathioprine as a single drug or in combination with steroids in the treatment of myasthenia gravis. *J. Neurol.* 1988;**235**(8):449-453.

6. La Mantia L, Mascoli N, Milanese C. Azathioprine. Safety profile in multiple sclerosis patients. *Neurol. Sci.* 2007;**28**(6):299-303.

7. Tiede I, Fritz G, Strand S, et al. CD28-dependent Rac1 activation is the molecular target of azathioprine in primary human CD4+ T lymphocytes. *J. Clin. Invest.* 2003;**111**(8):1133-1145.

8. Han HK, Amidon GL. Targeted prodrug design to optimize drug delivery. *A.A.P.S. PharmSci.* 2000;**2**(1):E6.

9. Lee AU, Farrell GC. Mechanism of azathioprine-induced injury to hepatocytes: roles of glutathione depletion and mitochondrial injury. *J. Hepatol.* 2001;**35**(6):756-764.
10. Coulthard S, Hogarth L. The thiopurines: an update. *Invest. New Drugs.* 2005;**23**(6):523-532.

11. Green CJ, Mee AS. Re-introduction of azathioprine in previously intolerant patients. *Eur. J. Gastroenterol. Hepatol.* 2006;**18**(1):17-19.

12. Bo J, Schrøder H, Kristinsson J, et al. Possible carcinogenic effect of 6-mercaptopurine on bone marrow stem cells: relation to thiopurine metabolism. *Cancer.* 1999;**86**(6):1080-1086.

13. de Boer NK, Mulder CJ, van Bodegraven AA. Myelotoxicity and hepatotoxicity during azathioprine therapy. *Neth. J. Med.* 2005;**63**(11):444-446.

14. Riccioni G, Bucciarelli V, Di Ilio E, et al. Recurrent atrial fibrillation in a patient with ulcerative colitis treated with azathioprine: case report and review of the literature. *Int. J. Immunopathol. Pharmacol.* 2011;**24**(1):247-249.

15. Manzano-Alonso ML, Castellano-Tortajada G. Reactivation of hepatitis B virus infection after cytotoxic chemotherapy or immunosuppressive therapy. *World. J. Gastroenterol.* 2011;**17**(12):1531-1537.

16. Lennard L. The clinical pharmacology of 6-mercaptopurine. *Eur. J. Clin. Pharmacol.* 1992;**43**(4):329-339.

17. Appell ML, Wennerstrand P, Peterson C, et al. Characterization of a novel sequence variant, TPMT*28, in the human thiopurine methyltransferase gene. *Pharmacogenet. Genomics.* 2010;**20**(11):700-707.

18. Fessing MY, Krynetski EY, Zambetti GP, et al. Functional characterization of the human thiopurine S-methyltransferase (TPMT) gene promoter. *Eur. J. Biochem.* 1998;**256**(3):510-517.

19. Spire-Vayron de la Moureyre C, Debuysère H, Fazio F, et al. Characterization of a variable number tandem repeat region in the thiopurine S-methyltransferase gene promoter. *Pharmacogenetics.* 1999;**9**(2):189-198.

20. Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. *Am. J. Hum. Genet.* 1980;**32**(5):651-662.

21. Hamdan-Khalil R, Gala JL, Allorge D, et al. Identification and functional analysis of two rare allelic variants of the thiopurine S-methyltransferase gene, TPMT*16 and TPMT*19. *Biochem. Pharmacol.* 2005;**69**(3):525-529.

22. Wang L, Nguyen TV, McLaughlin RW, et al. Human thiopurine S-methyltransferase pharmacogenetics: variant allozyme misfolding and aggresome formation. *Proc. Natl. Acad. Sci. U. S. A.* 2005;**102**(26):9394-9399.

23. Schaeffeler E, Lang T, Zanger UM, et al. High-throughput genotyping of thiopurine S-methyltransferase by denaturing HPLC. *Clin. Chem.* 2001;**47**(3):548-555.

24. Spire-Vayron de la Moureyre C, Debuysère H, Sabbagh N, et al. Detection of known and new mutations in the thiopurine S-methyltransferase gene by single-strand conformation polymorphism analysis. *Hum. Mutat.* 1998;**12**(3):177-185.

25. Tamm R, Oselin K, Kallassalu K, et al. Thiopurine S-methyltransferase (TPMT) pharmacogenetics: three new mutations and haplotype analysis in the Estonian population. *Clin. Chem. Lab. Med.* 2008;**46**(7):974-979.

26. Hamdan-Khalil R, Allorge D, Lo-Guidice JM, et al. In vitro characterization of four novel non-functional variants of the thiopurine S-methyltransferase. *Biochem. Biophys. Res. Commun.* 2003;**309**(4):1005-1010.

27. Hon YY, Fessing MY, Pui CH, et al. Polymorphism of the thiopurine S-methyltransferase gene in African-Americans. *Hum. Mol. Genet.* 1999;**8**(2):371-376.

28. Schaeffeler E, Fischer C, Brockmeier D, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. *Pharmacogenetics.* 2004;**14**(7):407-417.

29. Sasaki T, Goto E, Konno Y, et al. Three novel single nucleotide polymorphisms of the human thiopurine S-methyltransferase gene in Japanese individuals. *Drug. Metab. Pharmacokinet.* 2006;**21**(4):332-336.

30. Schaeffeler E, Eichelbaum M, Reinisch W, et al. Three novel thiopurine S-methyltransferase allelic variants (TPMT*20, *21, *22) - association with decreased enzyme function. *Hum. Mutat.* 2006;**27**(9):976.

31. Lindqvist M, Skoglund K, Karlgren A, et al. Explaining TPMT genotype/phenotype discrepancy by haplotyping of TPMT*3A and identification of a novel sequence variant, TPMT*23. *Pharmacogenet. Genomics.* 2007;**17**(10):891-895.

32. Garat A, Cauffiez C, Renault N, et al. Characterisation of novel defective thiopurine S-methyltransferase allelic variants. *Biochem. Pharmacol.* 2008;**76**(3):404-415.

33. Evans WE, Hon YY, Bomgaars L, et al. Preponderance of thiopurine S-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. *J. Clin. Oncol.* 2001;**19**(8):2293-2301.

34. Salavaggione OE, Wang L, Wiepert M, et al. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. *Pharmacogenet. Genomics.* 2005;**15**(11):801-815.

35. Krynetski EY, Schuetz JD, Galpin AJ, et al. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* 1995;**92**(4):949-953.

36. Wang L, Sullivan W, Toft D, et al. Thiopurine S-methyltransferase pharmacogenetics: chaperone protein association and allozyme degradation. *Pharmacogenetics.* 2003;**13**(9):555-564.

37. Weinshilboum R, Wang L. Pharmacogenetics: inherited variation in amino acid sequence and altered protein quantity. *Clin. Pharmacol. Ther.* 2004;**75**(4):253-258.

38. Otterness D, Szumlanski C, Lennard L, et al. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. *Clin. Pharmacol. Ther.* 1997;**62**(1):60-73.

39. Schaeffeler E, Stanulla M, Greil J, et al. A novel TPMT missense mutation associated with TPMT deficiency in a 5-year-old boy with ALL. *Leukemia.* 2003;**17**(7):1422-1424.

40. Lindqvist M, Haglund S, Almer S, et al. Identification of two novel sequence variants affecting thiopurine methyltransferase enzyme activity. *Pharmacogenetics.* 2004;**14**(4):261-265.

41. Kham SK, Soh CK, Aw DC, et al. TPMT*26 (208F-->L), a novel mutation detected in a Chinese. *Br. J. Clin. Pharmacol.* 2009;**68**(1):120-123.

42. Feng Q, Vannaprasaht S, Peng Y, et al. Thiopurine S-methyltransferase pharmacogenetics: functional characterization of a novel rapidly degraded variant allozyme. *Biochem. Pharmacol.* 2010;**79**(7):1053-1061.

43. Appell ML, Wennerstrand P, Peterson C, et al. Characterization of a novel sequence variant, TPMT*28, in the human thiopurine methyltransferase gene. *Pharmacogenet. Genomics.* 2010;**20**(11):700-707.

44. Evans WE, Horner M, Chu YQ, et al. Altered mercaptopurine metabolism, toxic effects, and dosage requirement in a thiopurine methyltransferase-deficient child with acute lymphocytic leukemia. *J. Pediatr.* 1991;**119**(6):985-989.

45. Lennard L, Davies HA, Lilleyman JS. Is 6-thioguanine more appropriate than 6-mercaptopurine for children with acute lymphoblastic leukaemia? *Br. J. Cancer.* 1993;68(1):186-190.

46. McLeod HL, Krynetski EY, Wilimas JA, et al. Higher activity of polymorphic thiopurine S-methyltransferase in erythrocytes from neonates compared to adults. *Pharmacogenetics.* 1995;5(5):281-286.

47. Kudo M, Moteki T, Sasaki T, et al. Functional characterization of human xanthine oxidase allelic variants. *Pharmacogenet. Genomics.* 2008;**18**(3):243-251.

48. Ansari A, Aslam Z, De Sica A, et al. Influence of xanthine oxidase on thiopurine metabolism in Crohn's disease. *Aliment. Pharmacol. Ther.* 2008;**28**(6):749-757.

49. Enroth C, Eger BT, Okamoto K, et al. Crystal structures of bovine milk xanthine dehydrogenase and xanthine oxidase: structure-based mechanism of conversion. *Proc. Natl. Acad. Sci. U. S. A.* 2000;**97**(20):10723-10728.

50. Pritsos CA. Cellular distribution, metabolism and regulation of the xanthine oxidoreductase enzyme system. *Chem. Biol. Interact.* 2000;**129**(1-2):195-208.

51. Jeroudi MO, Hartley CJ, Bolli R. Myocardial reperfusion injury: role of oxygen radicals and potential therapy with antioxidants. *Am. J. Cardiol.* 1994;**73**(6):2B-7B.

52. Németh I, Tálosi G, Papp A, et al. Xanthine oxidase activation in mild gestational hypertension. *Hypertens. Pregnancy.* 2002;**21**(1):1-11.

53. Guerciolini R, Szumlanski C, Weinshilboum RM. Human liver xanthine oxidase: nature and extent of individual variation. *Clin. Pharmacol. Ther.* 1991;**50**(6):663-672.

54. Saruwatari J, Nakagawa K, Shindo J, et al. A population phenotyping study of three drug-metabolizing enzymes in Kyushu, Japan, with use of the caffeine test. *Clin. Pharmacol. Ther.* 2002;**72**(2):200-208.

55. Carrillo JA, Benítez J. Caffeine metabolism in a healthy Spanish population: N-acetylator phenotype and oxidation pathways. *Clin. Pharmacol. Ther.* 1994;**55**(3):293-304.

56. Serre-Debeauvais F, Bayle F, Amirou M, et al. Hematotoxicity caused by azathioprine genetically determined and aggravated by xanthine oxidase deficiency in a patient following renal transplantation. *Presse. Med.* 1995;**24**(21):987-988.

57. Aklillu E, Carrillo JA, Makonnen E, et al. Xanthine oxidase activity is influenced by environmental factors in Ethiopians. *Eur. J. Clin. Pharmacol.* 2003;**59**(7):533-536.

58. Yamaguchi Y, Matsumura T, Ichida K, et al. Human xanthine oxidase changes its substrate specificity to aldehyde oxidase type upon mutation of amino acid residues in the active site: roles of active site residues in binding and activation of purine substrate. *J. Biochem.* 2007;**141**(4):513-524.

59. Ichida K, Amaya Y, Kamatani N, et al. Identification of two mutations in human xanthine dehydrogenase gene responsible for classical type I xanthinuria. *J. Clin. Invest.* 1997;**99**(10):2391-2397.

60. Levartovsky D, Lagziel A, Sperling O, et al. XDH gene mutation is the underlying cause of classical xanthinuria: a second report. *Kidney. Int.* 2000;**57**(6):2215-2220.

61. Sakamoto N, Yamamoto T, Moriwaki Y, et al. Identification of a new point mutation in the human xanthine dehydrogenase gene responsible for a case of classical type I xanthinuria. *Hum. Genet.* 2001;**108**(4):279-283.

62. Gok F, Ichida K, Topaloglu R. Mutational analysis of the xanthine dehydrogenase gene in a Turkish family with autosomal recessive classical xanthinuria. *Nephrol. Dial. Transplant.* 2003;**18**(11):2278-2283.

63. Yang J, Kamide K, Kokubo Y, et al. Associations of hypertension and its complications with variations in the xanthine dehydrogenase gene. *Hypertens. Res.* 2008;**31**(5):931-940.

64. Jurecka A, Stiburkova B, Krijt J, et al. Xanthine dehydrogenase deficiency with novel sequence variations presenting as rheumatoid arthritis in a 78-year-old patient. *J. Inherit. Metab. Dis.* 2010.

65. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA.* 1998;**279**(15):1200-1205.

66. Shastry BS. Pharmacogenetics and the concept of individualized medicine. *Pharmacogenomics. J.* 2006;**6**(1):16-21.

67. Eichelbaum M, Ingelman-Sundberg M, Evans WE. Pharmacogenomics and individualized drug therapy. *Annu. Rev. Med.* 2006;**57**:119-137.

68. Herrlinger KR, Jewell DP. Review article: interactions between genotype and response to therapy in inflammatory bowel diseases. *Aliment. Pharmacol. Ther.* 2006;**24**(10):1403-1412.

69. Goldstein DB, Need AC, Singh R, et al. Potential genetic causes of heterogeneity of treatment effects. *Am. J. Med.* 2007;**120**(4 Suppl 1):S21-25.

70. Kudo M, Sasaki T, Ishikawa M, et al. Kinetics of 6-thioxanthine metabolism by allelic variants of xanthine oxidase. *Drug. Metab. Pharmacokinet.* 2010;**25**(4):361-366.

71. Wong DR, Derijks LJ, den Dulk MO, et al. The role of xanthine oxidase in thiopurine metabolism: a case report. *Ther. Drug. Monit.* 2007;**29**(6):845-848.

72. Ansari A, Arenas M, Greenfield SM, et al. Prospective evaluation of the pharmacogenetics of azathioprine in the treatment of inflammatory bowel disease. *Aliment. Pharmacol. Ther.* 2008;**28**(8):973-983.

73. Wahed M, Louis-Auguste JR, Baxter LM, et al. Efficacy of methotrexate in Crohn's disease and ulcerative colitis patients unresponsive or intolerant to azathioprine /mercaptopurine. *Aliment. Pharmacol. Ther.* 2009;**30**(6):614-620.

74. Schymkowitz J, Borg J, Stricher F, et al. The FoldX web server: an online force field. *Nucleic. Acids. Res.* 2005;**3**:W382-388.

75. Van Durme J, Delgado J, Stricher F, et al. A graphical interface for the FoldX forcefield. *Bioinformatics.* 2011;**27**(12):1711-1712.

76. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005;**21**(2):263-265.

77. Schaid DJ, Rowland CM, Tines DE, et al. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.* 2002;**70**(2):425-434.

78. http://www.ncbi.nlm.nih.gov/SNP/ftplist2.cgi?fold=ftp://ftp.ncbi.nih.gov/snp/specs/Submission/

79. Suarez-Almazor ME, Spooner C, Belseck E. Azathioprine for treating rheumatoid arthritis. *Cochrane. Database. Syst. Rev.* 2000;**4**:CD001461.

80. Mertens HG, Hertel G, Reuther P, et al. Effect of immunosuppressive drugs (azathioprine). *Ann. N. Y. Acad. Sci.* 1981;**377**:691-699.

81. Mantegazza R, Antozzi C, Peluchetti D, et al. Azathioprine as a single drug or in combination with steroids in the treatment of myasthenia gravis. *J. Neurol.* 1988;**235**(8):449-453.

82. Cendrowski WS. Therapeutic trial of Imuran (azathioprine) in multiple sclerosis. *Acta. Neurol. Scand.* 1971;**47**(2):254-260.

83. Invernizzi P, Benedetti MD, Poli S, et al. Azathioprine in multiple sclerosis. *Mini. Rev. Med. Chem.* 2008;**8**(9):919-926.

84. Richman DP, Agius MA. Treatment of autoimmune myasthenia gravis. *Neurology.* 2003;**61**(12):1652-1661.

85. Martin LK, Werth V, Villanueva E, et al. Interventions for pemphigus vulgaris and pemphigus foliaceus. *Cochrane. Database. Syst. Rev.* 2009;**1**:CD006263.

86. Thiele DL. Autoimmune hepatitis. *Clin. Liver. Dis.* 2005;**9**(4):635-646

87. Weinshilboum RM. Pharmacogenomics: catechol O-methyltransferase to thiopurine S-methyltransferase. *Cell. Mol. Neurobiol.* 2006;**26**(4-6):539-561.

88. Ujiie S, Sasaki T, Mizugaki M, et al. Functional characterization of 23 allelic variants of thiopurine S-methyltransferase gene (TPMT*2 - *24). *Pharmacogenet. Genomics.* 2008;**18**(10):887-893.

89. Palaniswamy R, Teglund S, Lauth M, et al. Genetic variations regulate alternative splicing in the 5' untranslated regions of the mouse glioma-associated oncogene 1, Gli1. *B.M.C. Mol. Biol.* 2010;**11**:32.

90. Hull J, Campino S, Rowlands K, et al. Identification of common genetic variation that modulates alternative splicing. *PLoS. Genet.* 2007;**3**(6):e99.

91. Lamond AI, Konarska MM, Sharp PA. A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. *Genes. Dev.* 1987;**1**(6):532-543.

92. Sawa H, Shimura Y. Association of U6 snRNA with the 5'-splice site region of pre-mRNA in the spliceosome. *Genes. Dev.* 1992;**6**(2):244-254.

93. Santoro A, Cannella S, Trizzino A, et al. Mutations affecting mRNA splicing are the most common molecular defect in patients with familial hemophagocytic lymphohistiocytosis type 3. *Haematologica.* 2008;**93**(7):1086-1090.

94. Zhuang Y, Weiner AM. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes. Dev.* 1989;**3**(10):1545-1552.

95. Grainger RJ, Beggs JD. Prp8 protein: at the heart of the spliceosome. *RNA.* 2005;**11**(5):533-557.

96. Ameyaw MM, Collie-Duguid ES, Powrie RH, et al. Thiopurine methyltransferase alleles in British and Ghanaian populations. *Hum. Mol. Genet.* 1999;**8**(2):367-370.

97. Schaeffeler E, Zanger UM, Eichelbaum M, et al. Highly multiplexed genotyping of thiopurine s-methyltransferase variants using MALD-TOF mass spectrometry: reliable genotyping in different ethnic groups. *Clin. Chem.* 2008;**54**(10):1637-1647.

98. Cheon JH, Kim JH, Kim BY, et al. Allele frequency of thiopurine methyltransferase and inosine triphosphate pyrophosphatase gene polymorphisms in Korean patients with inflammatory bowel diseases. *Hepatogastroenterology.* 2009;**56**(90):421-423.

99. Smith MA, Marinaki AM, Arenas M, et al. Novel pharmacogenetic markers for treatment outcome in azathioprine-treated inflammatory bowel disease. *Aliment. Pharmacol. Ther.* 2009;**30**(4):375-384.

100. Spire-Vayron de la Moureyre C, Debuysere H, Mastain B, et al. Genotypic and phenotypic analysis of the polymorphic thiopurine S-methyltransferase gene (TPMT) in a European population. *Br. J. Pharmacol.* 1998;**125**(4):879-887.

101. McLeod HL, Pritchard SC, Githang'a J, et al. Ethnic differences in thiopurine methyltransferase pharmacogenetics: evidence for allele specificity in Caucasian and Kenyan individuals. *Pharmacogenetics.*1999;**9**(6):773-776.

102. Rossi AM, Bianchi M, Guarnieri C, et al. Genotype-phenotype correlation for thiopurine S-methyltransferase in healthy Italian subjects. *Eur. J. Clin. Pharmacol.* 2001;**57**(1):51-54.

103. Rochette J, Le Gac G, Lassoued K, et al. Factors influencing disease phenotype and penetrance in HFE haemochromatosis. *Hum. Genet.* 2010;**128**(3):233-248.

104. Meisler MH, O'Brien JE, Sharkey LM. Sodium channel gene family: epilepsy mutations, gene interactions and modifier effects. *J. Physiol.* 2010;**588**(Pt 11):1841-1848.

105. Sánchez IE, Beltrao P, Stricher F, et al. Genome-wide prediction of SH2 domain targets using structural information and the FoldX algorithm. *PLoS. Comput. Biol.* 2008;**4**(4):e1000052.

106. Strachan T, Read AP. Human Molecular Genetics. *2nd edition. New York: Wiley-Liss*; 1999. Chapter 1. DNA structure and gene expression. Bios Scientific Publishers, an imprint of Taylor & Francis Group

107. Gu S, Jin L, Zhang F, et al. Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat. Struct. Mol. Biol.* 2009;**16**(2):144-150.

108. Lennard L, Maddocks JL. Assay of 6-thioguanine nucleotide, a major metabolite of azathioprine, 6-mercaptopurine and 6-thioguanine, in human red blood cells. *J. Pharm. Pharmacol.* 1983;**35**(1):15-18.

109. Cara CJ, Pena AS, Sans M, et al. Reviewing the mechanism of action of thiopurine drugs: towards a new paradigm in clinical practice. *Med. Sci. Monit.* 2004;**10**(11):RA247-54.

APPENDIX 1

Summary of detected intronic and exonic *TPMT* SNPs[a]

P	AZA response	Int 3		Ex 5	Int 5	Int 6	Exon 7		Exon 10
		T140+ 114A	A141- 101T	C339T	T366+ 58C	G420- 4A	G460A	C474T	A719G
P1	I	wt	het	wt	het	wt	wt	het	wt
P2	I	wt	het	wt	het	wt	wt	het	wt
P3	I	wt	wt	wt	wt	wt	wt	wt	wt
P4	I	wt	wt	wt	wt	wt	wt	wt	wt
P5	I	wt	wt	wt	wt	wt	wt	het	wt
P6	I	wt	het	wt	het	wt	wt	het	wt
P7	I	wt	wt	wt	wt	wt	wt	wt	wt
P8	I	wt	hom	wt	hom	wt	wt	wt	wt
P9	I	het	hom	wt	hom	wt	het	wt	het
P10	I	wt	het	wt	het	wt	wt	wt	wt
P11	I	het	het	wt	het	wt	het	wt	het
P12	I	het	het	wt	het	wt	het	wt	het
P13	I	wt	het	wt	het	wt	wt	het	wt
P14	I	wt	het	wt	het	wt	wt	het	wt
P15	I	wt	het	wt	het	wt	wt	wt	wt
P16	I	wt	hom	wt	hom	wt	wt	wt	wt
P17	I	wt	het	wt	het	wt	wt	wt	wt
P18	I	wt	het	wt	het	het	wt	wt	wt
P19	I	het	het	wt	het	wt	het	wt	het
P20	I	wt	wt	wt	wt	wt	wt	wt	wt
P21	I	wt	wt	wt	wt	wt	wt	het	wt
P22	I	wt	het	wt	het	wt	wt	wt	wt
P23	I	wt	wt	wt	wt	wt	wt	wt	wt
P24	I	wt	wt	wt	wt	wt	wt	hom	wt
P25	I	wt	het	wt	het	wt	wt	wt	wt
P26	U	wt	het	wt	het	wt	wt	wt	wt
P27	U	wt	wt	wt	het	wt	wt	hom	wt
P28	U	wt	hom	wt	hom	wt	het	wt	het
P29	U	wt	het	wt	het	wt	wt	het	wt
P30	U	wt	wt	wt	wt	wt	wt	het	wt
P31	U	wt	hom	het	hom	wt	wt	het	wt
P32	U	wt	het	wt	het	wt	wt	wt	wt
P33	U	wt	het	wt	het	wt	wt	het	wt
P34	U	wt	hom	wt	hom	wt	wt	wt	wt
P35	U	wt	het	wt	het	wt	wt	hom	wt
P36	U	wt	wt	wt	wt	wt	wt	wt	wt
P37	U	wt	hom	wt	hom	wt	wt	wt	wt
P38	U	wt	het	wt	het	wt	wt	wt	wt
P39	U	wt	wt	wt	wt	wt	wt	wt	wt
P40	U	wt	wt	wt	wt	wt	wt	wt	wt
P41	U	wt	hom	wt	hom	wt	wt	wt	wt
P42	R	wt	wt	wt	wt	wt	wt	wt	wt
P43	R	wt	hom	wt	hom	wt	het	wt	het
P44	R	wt	wt	wt	wt	wt	wt	wt	wt
P45	R	wt	hom	wt	hom	wt	wt	wt	wt
P46	R	wt	het	wt	het	wt	wt	het	wt
P47	R	wt	het	wt	het	wt	wt	het	wt
P48	R	wt	het	wt	het	wt	wt	het	wt

P	AZA response	Int 3		Ex 5	Int 5	Int 6	Exon 7		Exon 10
		T140+114A	A141-101T	C339T	T366+58C	G420-4A	G460A	C474T	A719G
P49	R	wt	het	wt	het	wt	wt	wt	wt
P50	R	wt	het	wt	het	wt	wt	hom	wt
P51	R	wt	het	wt	het	wt	wt	het	wt
P52	R	wt	het	wt	het	wt	wt	het	wt
P53	R	wt	hom	wt	hom	wt	wt	het	wt
P54	R	wt	hom	wt	hom	wt	wt	het	wt
P55	R	wt	hom	wt	hom	wt	wt	het	wt
P56	R	wt	het	wt	het	wt	wt	wt	wt
P57	R	wt	hom	wt	hom	wt	wt	wt	wt
P58	R	wt	wt	wt	wt	wt	wt	wt	wt
P59	R	wt	hom	wt	hom	wt	wt	het	wt
P60	R	wt	het	wt	het	wt	wt	wt	wt
P61	R	wt	het	wt	het	wt	wt	wt	wt
P62	R	wt	wt	wt	wt	wt	wt	wt	wt
P63	R	wt	hom	wt	hom	wt	wt	wt	wt
P64	R	wt	het	het	het	wt	wt	het	wt
P65	R	wt	het	wt	hom	wt	wt	het	wt
P66	R	wt	wt	wt	wt	wt	wt	wt	wt
P67	R	wt	het	wt	het	wt	wt	wt	wt
P68	R	wt	het	wt	het	wt	wt	het	wt
P69	R	wt	wt	wt	wt	wt	wt	wt	wt
P70	R	wt	het	wt	het	wt	wt	wt	wt
P71	R	wt	hom	wt	hom	wt	wt	wt	wt

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Int: intron; Ex: exon; het: heterozygous; hom: homozygous; wt: wild type.

APPENDIX 2

Summary of detected exonic *XDH* SNPs (Exon 7 - Exon 13)[a]

P	AZA response	Ex 7	Ex 8	Ex 10			Ex 11	Ex 13
		G514A	G627A	G822A	C837T	C859G	G1004A	G1194A
P1	I	wt	wt	wt	wt	wt	wt	wt
P2	I	wt	wt	wt	wt	wt	wt	wt
P3	I	wt	wt	wt	wt	wt	wt	wt
P4	I	wt	wt	wt	wt	wt	wt	wt
P5	I	wt	wt	wt	wt	wt	wt	wt
P6	I	wt	wt	wt	wt	wt	wt	wt
P7	I	wt	wt	wt	wt	wt	wt	wt
P8	I	wt	wt	wt	wt	wt	wt	wt
P9	I	wt	wt	wt	wt	wt	wt	wt
P10	I	het	het	wt	het	wt	wt	wt
P11	I	wt	wt	wt	wt	het	wt	wt
P12	I	het	het	wt	het	wt	wt	wt
P13	I	wt	wt	wt	wt	wt	wt	wt
P14	I	wt	wt	wt	wt	wt	wt	wt
P15	I	wt	wt	wt	wt	wt	wt	wt
P16	I	wt	wt	wt	wt	wt	wt	wt
P17	I	wt	wt	wt	wt	wt	wt	wt
P18	I	wt	wt	wt	wt	wt	wt	wt
P19	I	wt	wt	wt	wt	wt	wt	wt
P20	I	wt	wt	wt	wt	wt	wt	het
P21	I	wt	wt	wt	wt	wt	het	wt
P22	I	wt	wt	wt	wt	wt	wt	wt
P23	I	wt	wt	wt	wt	wt	wt	wt
P24	I	wt	wt	wt	wt	wt	wt	wt
P25	I	wt	wt	wt	wt	wt	wt	wt
P26	U	wt	wt	wt	wt	wt	wt	wt
P27	U	wt	wt	wt	wt	wt	wt	wt
P28	U	wt	wt	wt	wt	wt	wt	wt
P29	U	wt	wt	wt	wt	wt	wt	wt
P30	U	wt	wt	wt	wt	wt	wt	wt
P31	U	wt	wt	wt	wt	wt	wt	wt
P32	U	wt	wt	wt	wt	wt	wt	wt
P33	U	wt	wt	wt	wt	wt	wt	wt
P34	U	wt	wt	wt	wt	wt	wt	wt
P35	U	het	wt	wt	het	wt	wt	wt
P36	U	het	het	wt	het	wt	wt	wt
P37	U	wt	wt	wt	wt	wt	wt	wt
P38	U	wt	wt	wt	wt	wt	wt	wt
P39	U	wt	wt	wt	wt	wt	wt	wt
P40	U	wt	wt	wt	wt	wt	wt	wt
P41	U	wt	wt	wt	wt	wt	wt	wt
P42	R	wt	wt	wt	wt	wt	wt	wt
P43	R	wt	wt	wt	wt	wt	wt	wt
P44	R	wt	wt	wt	wt	wt	wt	wt
P45	R	wt	wt	wt	wt	wt	wt	wt
P46	R	wt	wt	wt	wt	wt	wt	wt
P47	R	wt	wt	wt	wt	wt	wt	wt
P48	R	wt	wt	wt	wt	wt	wt	wt
P49	R	wt	wt	wt	wt	wt	wt	wt

P	AZA response	Ex 7	Ex 8	Ex 10			Ex 11	Ex 13
		G514A	G627A	G822A	C837T	C859G	G1004A	G1194A
P50	R	wt	wt	wt	wt	wt	wt	wt
P51	R	wt	wt	wt	het	wt	wt	wt
P52	R	wt	wt	wt	wt	wt	wt	wt
P53	R	wt	wt	wt	wt	wt	wt	wt
P54	R	wt	wt	wt	wt	wt	wt	wt
P55	R	wt	wt	wt	wt	wt	wt	wt
P56	R	het	het	wt	het	wt	wt	wt
P57	R	het	wt	wt	het	wt	wt	wt
P58	R	wt	wt	wt	wt	wt	wt	wt
P59	R	het	het	het	wt	wt	wt	wt
P60	R	wt	wt	wt	wt	wt	wt	wt
P61	R	het	het	wt	het	wt	wt	wt
P62	R	wt	wt	wt	wt	wt	wt	wt
P63	R	wt	wt	wt	wt	wt	wt	wt
P64	R	wt	wt	wt	wt	wt	wt	wt
P65	R	het	het	wt	het	wt	wt	wt
P66	R	het	het	wt	het	wt	wt	wt
P67	R	wt	wt	wt	wt	wt	wt	wt
P68	R	wt	wt	wt	wt	wt	wt	wt
P69	R	wt	wt	wt	wt	wt	wt	wt
P70	R	wt	wt	wt	wt	wt	wt	wt
P71	R	het	het	wt	wt	wt	wt	wt

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Ex: exon; het: heterozygous; hom: homozygous; wt: wild type.

Summary of detected exonic *XDH* SNPs (Exon 14 - Exon 21)[a]

P	AZA Response	Ex 14		Ex 15	Ex 16	Ex 18	Ex 20	Ex 21
		C1274G	G1329A	C1509G	C1667T	A1936G	A2107G	C2211T
P1	I	wt	wt	wt	wt	wt	wt	wt
P2	I	wt	wt	wt	het	het	het	het
P3	I	wt	wt	wt	wt	wt	wt	het
P4	I	wt	wt	wt	wt	wt	wt	wt
P5	I	wt	wt	wt	wt	wt	wt	wt
P6	I	wt	wt	wt	wt	wt	wt	wt
P7	I	wt	wt	wt	wt	wt	wt	hom
P8	I	wt	wt	wt	wt	wt	wt	wt
P9	I	wt	wt	wt	wt	wt	wt	het
P10	I	wt	wt	wt	wt	wt	wt	wt
P11	I	wt	wt	wt	wt	wt	wt	wt
P12	I	wt	wt	wt	wt	wt	wt	wt
P13	I	wt	wt	wt	wt	wt	wt	wt
P14	I	wt	wt	wt	wt	wt	wt	het
P15	I	wt	wt	wt	wt	wt	wt	hom
P16	I	wt	wt	wt	wt	wt	wt	het
P17	I	wt	wt	wt	wt	wt	wt	wt
P18	I	wt	wt	wt	wt	wt	wt	wt
P19	I	wt	wt	wt	wt	wt	wt	wt
P20	I	wt	wt	wt	wt	wt	wt	het
P21	I	wt	wt	wt	wt	wt	wt	het
P22	I	wt	wt	wt	wt	het	het	wt
P23	I	wt	wt	wt	wt	wt	wt	wt
P24	I	wt	wt	wt	wt	wt	wt	het

P	AZA Response	Ex 14		Ex 15	Ex 16	Ex 18	Ex 20	Ex 21
		C1274G	G1329A	C1509G	C1667T	A1936G	A2107G	C2211T
P25	I	wt	wt	wt	wt	wt	wt	wt
P26	U	wt	wt	wt	wt	wt	wt	wt
P26	U	wt	wt	wt	wt	wt	wt	wt
P27	U	wt	wt	het	wt	wt	wt	wt
P28	U	wt	wt	wt	wt	wt	wt	wt
P29	U	wt	wt	wt	wt	wt	wt	wt
P30	U	wt	wt	wt	wt	het	het	wt
P31	U	het	wt	wt	wt	het	het	wt
P32	U	wt	wt	wt	wt	wt	wt	wt
P33	U	wt	wt	wt	wt	wt	wt	het
P34	U	wt	wt	wt	wt	wt	wt	wt
P35	U	wt	wt	wt	wt	wt	wt	wt
P36	U	wt	wt	wt	wt	wt	wt	het
P37	U	wt	wt	wt	wt	wt	wt	het
P38	U	wt	wt	wt	het	wt	wt	wt
P39	U	wt	wt	wt	wt	wt	wt	wt
P40	U	wt	wt	wt	wt	het	het	wt
P41	U	wt	wt	wt	wt	het	het	wt
P42	R	wt	wt	wt	wt	wt	wt	wt
P43	R	wt	wt	wt	wt	wt	wt	wt
P44	R	wt	wt	wt	wt	wt	wt	het
P45	R	wt	wt	wt	wt	wt	wt	wt
P46	R	wt	wt	wt	wt	wt	wt	wt
P47	R	wt	wt	wt	wt	het	het	het
P48	R	wt	wt	wt	wt	wt	wt	wt
P49	R	wt	wt	wt	wt	wt	wt	wt
P50	R	wt	wt	wt	wt	wt	wt	hom
P51	R	wt	wt	wt	wt	wt	wt	hom
P52	R	wt	wt	wt	wt	wt	wt	het
P53	R	wt	wt	wt	wt	wt	wt	wt
P54	R	wt	wt	wt	wt	wt	wt	het
P55	R	wt	wt	wt	het	wt	wt	wt
P56	R	wt	wt	wt	wt	wt	wt	wt
P57	R	wt	wt	wt	wt	wt	wt	wt
P58	R	wt	wt	wt	wt	wt	wt	het
P59	R	wt	wt	wt	wt	wt	wt	het
P60	R	wt	wt	wt	wt	wt	wt	wt
P61	R	wt	het	wt	wt	wt	wt	wt
P62	R	wt	wt	wt	wt	wt	wt	wt
P63	R	wt	wt	wt	wt	wt	wt	wt
P64	R	wt	wt	wt	het	wt	wt	wt
P65	R	wt	wt	wt	wt	wt	wt	wt
P66	R	wt	wt	wt	wt	wt	wt	wt
P67	R	wt	wt	wt	wt	wt	wt	het
P68	R	wt	het	wt	wt	wt	wt	wt
P69	R	wt	wt	wt	wt	wt	wt	het
P70	R	wt	wt	wt	wt	wt	wt	wt
P71	R	wt	wt	wt	wt	wt	wt	hom

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Ex: exon; het: heterozygous; hom: homozygous; wt: wild type.

Summary of detected exonic *XDH* SNPs (Exon 26 - Exon 34)[a]

P	AZA Response	Ex 26 C2891T	Ex 27 T3030C	Ex 32 T3488C	Ex 33 T3536C	Ex 34 G3717A
P1	I	wt	hom	wt	wt	wt
P2	I	wt	het	wt	wt	wt
P3	I	wt	hom	wt	wt	wt
P4	I	wt	het	wt	wt	wt
P5	I	wt	het	wt	wt	wt
P6	I	wt	hom	wt	wt	wt
P7	I	wt	hom	wt	wt	wt
P8	I	wt	het	wt	wt	wt
P9	I	wt	hom	wt	wt	wt
P10	I	wt	hom	wt	wt	wt
P11	I	wt	het	wt	wt	het
P12	I	wt	het	wt	wt	wt
P13	I	wt	hom	wt	wt	wt
P14	I	wt	hom	wt	wt	wt
P15	I	wt	hom	wt	wt	wt
P16	I	wt	hom	wt	wt	wt
P17	I	wt	hom	wt	wt	wt
P18	I	wt	hom	wt	wt	wt
P19	I	wt	wt	wt	wt	wt
P20	I	wt	hom	wt	wt	wt
P21	I	wt	hom	wt	wt	wt
P22	I	wt	wt	wt	wt	wt
P23	I	wt	het	wt	wt	wt
P24	I	wt	hom	wt	wt	wt
P25	I	wt	het	wt	wt	wt
P26	U	wt	hom	wt	wt	wt
P27	U	wt	het	wt	wt	wt
P28	U	wt	het	wt	wt	wt
P29	U	wt	het	wt	wt	wt
P30	U	wt	wt	wt	wt	wt
P31	U	wt	het	wt	wt	wt
P32	U	wt	hom	wt	wt	wt
P33	U	wt	hom	wt	wt	wt
P34	U	wt	hom	wt	wt	wt
P35	U	wt	hom	wt	het	wt
P36	U	wt	hom	wt	wt	wt
P37	U	wt	hom	wt	wt	wt
P38	U	wt	het	wt	wt	wt
P39	U	wt	het	wt	wt	wt
P40	U	wt	het	wt	wt	wt
P41	U	wt	het	wt	wt	wt
P42	R	wt	hom	wt	wt	wt
P43	R	wt	het	wt	wt	wt
P44	R	wt	het	wt	wt	wt
P45	R	wt	wt	wt	wt	wt
P46	R	wt	het	wt	wt	wt
P47	R	wt	het	wt	wt	wt
P48	R	wt	het	wt	wt	wt
P49	R	wt	het	wt	wt	wt
P50	R	het	hom	wt	wt	wt
P51	R	wt	hom	wt	wt	wt
P52	R	wt	het	het	wt	wt

P	AZA Response	Ex 26 C2891T	Ex 27 T3030C	Ex 32 T3488C	Ex 33 T3536C	Ex 34 G3717A
P53	R	wt	het	wt	wt	wt
P54	R	wt	hom	wt	wt	wt
P55	R	wt	hom	wt	wt	wt
P56	R	wt	hom	wt	wt	wt
P57	R	wt	hom	wt	wt	wt
P58	R	wt	het	wt	wt	wt
P59	R	wt	het	wt	wt	wt
P60	R	wt	hom	wt	wt	wt
P61	R	wt	hom	wt	wt	wt
P62	R	wt	hom	wt	wt	wt
P63	R	wt	het	wt	wt	wt
P64	R	wt	hom	wt	wt	wt
P65	R	wt	hom	wt	wt	wt
P66	R	wt	het	wt	wt	wt
P67	R	wt	hom	wt	wt	wt
P68	R	wt	het	wt	wt	het
P69	R	wt	hom	wt	wt	wt
P70	R	wt	wt	wt	wt	wt
P71	R	wt	hom	wt	wt	wt

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Ex: exon; het: heterozygous; hom: homozygous; wt: wild type.

APPENDIX 3

Summary of detected intronic *XDH* SNPs (Intron 1 – Intron 23)[a]

P	AZA Response	Int 1 C100+ 27T	Int 4 A307- 33G	Int 8 T652- 21A	Int 17 G185 7-4A	Int 20 G2197+ 42C	Int 20 G2197+ 68A	Int 23 A2544+ 35T
P1	I	wt	wt	wt	wt	wt	wt	hom
P2	I	het	het	wt	wt	het	wt	hom
P3	I	wt	wt	wt	wt	het	het	hom
P4	I	wt	wt	wt	wt	wt	het	het
P5	I	wt	wt	wt	wt	wt	het	het
P6	I	wt	wt	wt	wt	het	het	hom
P7	I	wt	wt	wt	wt	hom	wt	hom
P8	I	wt	wt	wt	wt	wt	wt	wt
P9	I	wt	wt	wt	wt	het	het	hom
P10	I	wt	wt	wt	wt	wt	het	hom
P11	I	het	wt	wt	wt	wt	wt	het
P12	I	wt	wt	wt	wt	wt	wt	het
P13	I	wt	het	wt	wt	wt	hom	hom
P14	I	wt	wt	wt	wt	het	het	hom
P15	I	wt	wt	wt	wt	hom	wt	hom
P16	I	wt	wt	wt	wt	het	het	hom
P17	I	wt	het	wt	wt	wt	hom	hom
P18	I	wt	wt	wt	wt	wt	hom	hom
P19	I	wt	wt	het	wt	wt	wt	hom
P20	I	wt	het	wt	wt	het	wt	het
P21	I	wt	het	wt	wt	het	het	hom
P22	I	het	wt	wt	wt	wt	het	hom
P23	I	wt	wt	wt	wt	wt	het	wt
P24	I	wt	wt	wt	wt	het	het	hom
P25	I	wt	wt	wt	wt	wt	het	het
P26	U	wt	het	wt	wt	het	het	wt
P27	U	wt	wt	wt	wt	wt	wt	het
P28	U	wt	wt	wt	wt	wt	wt	het
P29	U	wt	wt	wt	wt	wt	het	het
P30	U	wt	het	wt	wt	wt	wt	het
P31	U	het	het	wt	wt	wt	wt	hom
P32	U	wt	wt	wt	wt	wt	hom	hom
P33	U	wt	wt	wt	wt	het	het	hom
P34	U	wt	wt	wt	wt	wt	hom	hom
P35	U	wt	wt	wt	wt	wt	het	hom
P36	U	wt	wt	wt	wt	het	wt	hom
P37	U	wt	het	wt	wt	het	het	hom
P38	U	wt	wt	wt	wt	wt	het	het
P39	U	wt	wt	wt	wt	wt	het	het
P40	U	wt	het	wt	wt	wt	het	hom
P41	U	wt	het	wt	het	wt	het	hom
P42	R	wt	wt	wt	wt	wt	hom	hom
P43	R	wt	wt	wt	wt	wt	wt	het
P44	R	het	wt	wt	wt	het	wt	het
P45	R	wt	wt	wt	wt	wt	wt	wt
P46	R	wt	wt	wt	wt	het	wt	het
P47	R	wt	wt	wt	wt	het	wt	hom
P48	R	wt	wt	wt	wt	wt	het	het

P	AZA Response	Int 1 C100+ 27T	Int 4 A307- 33G	Int 8 T652- 21A	Int 17 G185 7-4A	Int 20 G2197+ 42C	Int 20 G2197+ 68A	Int 23 C100+2 7T
P49	R	wt	wt	wt	wt	wt	het	wt
P50	R	wt	het	wt	wt	hom	wt	hom
P51	R	wt	het	wt	wt	hom	wt	hom
P52	R	wt	wt	wt	wt	het	wt	het
P53	R	wt	het	wt	wt	wt	wt	wt
P54	R	wt	wt	wt	wt	het	het	hom
P55	R	wt	wt	wt	wt	wt	hom	hom
P56	R	wt	het	wt	wt	het	wt	hom
P57	R	wt	wt	wt	wt	wt	het	hom
P58	R	wt	wt	wt	wt	het	wt	het
P59	R	wt	wt	wt	wt	het	wt	het
P60	R	wt	het	wt	wt	wt	het	het
P61	R	wt	wt	wt	wt	wt	het	hom
P62	R	wt	wt	wt	wt	het	het	hom
P63	R	wt	wt	wt	wt	wt	het	het
P64	R	wt	het	wt	wt	wt	wt	het
P65	R	wt	het	wt	wt	hom	wt	het
P66	R	wt	wt	wt	wt	wt	wt	wt
P67	R	wt	wt	wt	wt	het	het	hom
P68	R	wt	wt	wt	wt	wt	het	het
P69	R	wt	hom	wt	wt	het	wt	het
P70	R	wt	wt	wt	wt	wt	wt	wt
P71	R	wt	wt	wt	wt	hom	wt	hom

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Int: intron; het: heterozygous; hom: homozygous; wt: wild type.

Summary of detected intronic *XDH* SNPs (Intron 25 – 3'UTR)[a]

P	AZA Response	Int 25 G2824 -35A	Int 27 C3052- 12G	Int 30 A3352- 30C	Int 34 C3774+ 32G	3'UTR A4002+ 518G	3'UTR A4002+ 1477G	3'UTR C4002+ 1516T	3'UTR G4002+ 1555T
P1	I	wt	hom	wt	wt	het	het	het	wt
P2	I	wt	het	wt	het	het	wt	het	wt
P3	I	wt	hom	wt	wt	hom	wt	hom	wt
P4	I	wt	het	wt	wt	hom	wt	hom	wt
P5	I	wt	het	wt	wt	het	het	het	het
P6	I	wt	hom	wt	wt	het	het	het	het
P7	I	wt	hom	wt	wt	het	wt	het	wt
P8	I	wt	wt	wt	wt	wt	het	wt	het
P9	I	wt	hom	wt	wt	het	wt	het	wt
P10	I	wt	hom	het	wt	het	het	het	wt
P11	I	wt	het	wt	wt	wt	hom	wt	hom
P12	I	wt	het	wt	wt	hom	wt	hom	wt
P13	I	wt	hom	wt	wt	wt	het	het	het
P14	I	wt	hom	wt	wt	hom	wt	hom	wt
P15	I	wt	hom	wt	wt	wt	wt	wt	wt
P16	I	wt	hom	wt	wt	het	wt	het	wt
P17	I	wt	hom	wt	wt	het	het	het	het
P18	I	wt	hom	wt	wt	hom	wt	hom	wt
P19	I	wt	wt	wt	wt	wt	het	wt	het
P20	I	wt	hom	wt	wt	hom	wt	hom	wt
P21	I	wt	hom	wt	wt	hom	wt	hom	wt

P	AZA Response	Int 25 G2824 -35A	Int 27 C3052- 12G	Int 30 A3352- 30C	Int 34 C3774+ 32G	3'UTR			
						A4002+ 518G	A4002+ 1477G	C4002+ 1516T	G4002+ 1555T
P22	I	wt	het	wt	wt	hom	wt	hom	wt
P23	I	wt	het	wt	wt	hom	wt	hom	wt
P24	I	wt	hom	wt	wt	hom	het	het	het
P25	I	wt	wt	wt	wt	het	het	het	het
P26	U	wt	hom	wt	wt	het	het	het	het
P27	U	wt	het	wt	wt	wt	hom	wt	wt
P28	U	wt	het	wt	wt	wt	het	wt	hom
P29	U	wt	het	wt	wt	het	wt	het	wt
P30	U	wt	wt	wt	wt	wt	het	wt	het
P31	U	wt	het	wt	wt	wt	wt	wt	wt
P32	U	wt	hom	wt	wt	het	het	het	het
P33	U	wt	hom	het	wt	wt	het	wt	wt
P34	U	wt	hom	wt	wt	het	het	het	het
P35	U	wt	hom	wt	wt	wt	hom	wt	het
P36	U	wt	hom	het	wt	het	het	het	wt
P37	U	wt	hom	wt	wt	het	wt	het	wt
P38	U	wt	het	wt	wt	het	het	het	het
P39	U	wt	het	wt	wt	hom	wt	hom	wt
P40	U	wt	het	wt	wt	wt	wt	wt	wt
P41	U	wt	het	wt	het	het	wt	het	wt
P42	R	wt	hom	wt	wt	wt	het	wt	het
P43	R	wt	het	wt	wt	het	het	wt	hom
P44	R	wt	het	wt	wt	het	het	het	het
P45	R	wt	wt	wt	wt	wt	het	wt	het
P46	R	wt	het	wt	wt	wt	wt	wt	wt
P47	R	wt	het	wt	wt	hom	wt	hom	wt
P48	R	wt	het	hom	wt	het	het	het	wt
P49	R	wt	het	wt	wt	hom	wt	hom	wt
P50	R	wt	hom	wt	wt	het	wt	wt	wt
P51	R	wt	hom	wt	wt	het	het	het	het
P52	R	wt	het	wt	wt	het	het	het	het
P53	R	wt	het	wt	wt	het	wt	het	wt
P54	R	wt	hom	wt	wt	het	wt	het	wt
P55	R	wt	hom	wt	wt	wt	hom	wt	hom
P56	R	wt	hom	wt	wt	het	het	het	wt
P57	R	wt	hom	wt	wt	wt	het	wt	het
P58	R	wt	het	wt	wt	hom	wt	hom	wt
P59	R	wt	het	wt	wt	wt	hom	wt	hom
P60	R	wt	hom	het	wt	het	het	het	wt
P61	R	wt	hom	hom	wt	wt	hom	wt	het
P62	R	wt	hom	wt	het	het	wt	het	wt
P63	R	wt	het	wt	wt	het	het	het	het
P64	R	het	hom	het	wt	wt	het	wt	wt
P65	R	het	hom	wt	wt	wt	het	wt	wt
P66	R	wt	hom	wt	wt	wt	wt	wt	wt
P67	R	wt	hom	wt	wt	het	het	het	het
P68	R	wt	het	wt	wt	wt	wt	wt	hom
P69	R	wt	hom	wt	wt	hom	wt	hom	wt
P70	R	wt	wt	wt	wt	het	wt	wt	wt
P71	R	wt	hom	wt	wt	hom	wt	hom	wt

[a] Only SNPs that were detected in at least one patient are shown; P: patient; Int: intron; het: heterozygous; hom: homozygous; wt: wild type.

ACKNOWLEDGMENTS

Al termine di questo nuovo viaggio le persone da ringraziare sono veramente numerose.
I ringraziamenti più grandi vanno ovviamente a mia mamma e a Daniele, che, come sempre, mi sono stati vicini incoraggiandomi e sostenendomi nei momenti di sconforto. Senza di loro credo che questo percorso sarebbe stato molto più faticoso e pesante.
Un secondo ringraziamento doveroso va alla dottoressa Pia Bernasconi e a tutta l'Unità Operativa Neuromuscolare IV, Divisione Malattie Neuromuscolari della fondazione IRCCS "Carlo Besta", che mi hanno dato la possibilità di effettuare lo studio, credendo nella bontà del progetto, e mi hanno consigliato quando non vedevo soluzione ai problemi che si sono presentati in questi tre anni di lavoro.
Devo ringraziare di cuore Paola, Stefania e Cristina colleghe di "banco" per diversi anni con cui mi sono spesso confrontata e che mi hanno sempre dato validi consigli, ma soprattutto, amiche con cui ridere, scherzare e passare allegri pomeriggi anche al di fuori del laboratorio. Un grande ringraziamento anche alle new entry Claudia, Tiziana, Eduardo, Riccardo, Elisa e Simona con cui non lavoro a stretto contatto ma che sono persone molto allegre e disponibili. Non posso dimenticare Eleonora e Raffaella, nuove colleghe di "banco" da un anno a questa parte, sempre disponibili nell'aiutare e nel fare qualche sana risata.
Devo essere infinitamente grata, inoltre, a Dimos, insuperabile bioinformatico che mi ha dedicato tantissimo tempo, non solo per analizzare i miei dati a livello statistico, ma anche per spiegarmi, in modo chiaro e semplice, una piccola parte del mondo della bioinformatica. Inoltre devo ringraziare il dottor Lorenzo Maggi e la dottoressa Giorgia Camera che hanno reperito e letto le cartelle cliniche dei pazienti per ottenere i dati clinici da inserire in questo lavoro.
Un grazie di cuore, inoltre, va a tutti gli amici che sono con me fin dall'infanzia e ai compagni di avventure della Croce Bianca di Bergamo, che mi hanno sopportato, confortato e divertito. Tra loro devo ricordare Walter e Ale, Rita e Mario, Alice e Fabio.
Grazie ancora a tutti, spero di non essermi dimenticata di nessuno, se così non fosse sappiate che, anche se non vi ho nominato, siete comunque nel mio cuore.