Aus dem Institut für Pharmakologie und Toxikologie
Direktor: Prof. Dr. T. Gudermann
des Fachbereichs Medizin der Philipps-Universität Marburg

Human arsenic methyltransferase pharmacogenetics:

functional studies of common polymorphisms

and its impact on medicine

Inaugural-Dissertation zur
Erlangung des Doktorgrades
der gesamten Humanmedizin

dem
Fachbereich Medizin
der Philipps-Universität Marburg
vorgelegt von

Annette Friederike Klumpp
aus Herne

Marburg, 2008

Angenommen vom Fachbereich Medizin

der Philipps-Universität Marburg am: 16.05.2008

Gedruckt mit der Genehmigung des Fachbereichs.

Dekan: Prof. Dr. M. Rothmund

Referent: Prof. Dr. Dr. J. Krieglstein

Korreferent: PD Dr. S. Müller-Brüsselbach

"....Si toutes ces excuses ne suffisent pas,
je vieux bien dédier ce livre à l'enfant
qu'a été autrefois cette grande personne.
Toutes les grandes personnes ont d'abord été des
enfants. (Mais peu d'entre s'en souviennent)...."

Antoine de Saint-Exupéry
Le Petit Prince

Table of Contents

List of Figures

List of Tables

Abbreviations

A	Adenosine
AdoMet	S-adenosyl-L-methionine
Arg	Arginine
AS3MT	Arsenic methyltransferase
As^{III}	Trivalent arsenite
As^{V}	Pentavalent arsenate
ATP	Adenosine triphosphate
bp	Base pairs
C	Cytidine
°C	Temperature in Celsius
^{14}C	Carbon 14
cDNA	Complementary deoxynucleic acid
CH_3CH_2OH	Ethanol
$CH_3CHOHCH_3$	Isopropanol
$CHCl_3$	Chloroform
Ci	Curie
CO_2	Carbon dioxide
COS-1	African green monkey kidney cell line
CpG	Cytodine-phosphatidyl-guanosine
cpm	Counts per minute
cSNP	Coding single nucleotide polymorphism
DMA	Dimethylarsenic
DMEM	Dulbeccos's Modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxynucleic acid
dNTP	Deoxyribonucleotide triphosphate
E. coli	Escherichia coli
EDTA	Ethylenediamine tetraacetic acid
F	Forward
FBS	Fetal bovine serum

FR	Flanking region
g	Gram
g	Gravity
G	Guanosine
GSH	Glutathion
^{3}H	Tritium
HCl	Hydrogen chloride
HEK293	Human embryotic kidney cell line
HepG2	Human liver cell line
H_2O	Water
HSS	High speed supernatant
Ile	Isoleucine
$KC_2H_3O_2$	Potassium acetate
KCl	Potassium chloride
$K_2Cr_2O_7$	Potassium dichromate
KH_2PO_4	Potassium dihydrogen phosphate
KI	Potassium iodide
l	Liter
LAR	Luciferase assay reagent
µ	Micro-
MAT	Methionine adenosyltransferase
α-MEM	Modified Eagel medium
Met	Methionine
mg	Milligram
$MgCl_2$	Magnesiumchloride
min	Minute
mM	Millimolar
MMA	Monomethylarsenic
MOPS	3-N-morpholino-propanesulphonic-acid
6-MP	6-Mercaptopurine
mRNA	Messenger ribonucleic acid
N	Number
$NaAsO_2$	Sodium arsenite
NaCl	Sodium chloride

Na$_2$HPO$_4$	Disodium hydrogen phosphate
NaI	Sodium iodide
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
NRC	National Research Council
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
R	Reverse
RBC	Red blood cell
rpm	Rotations per minute
SAM	S-adenosyl-L-methionine (AdoMet)
SDS	Sodium dodecyl sulfate
sec	Second
SNP	Single nucleotide polymorphism
T	Thymidine
TAE	Tris-acetate-EDTA buffer
Thr	Threonine
Tm	Melting temperature
TpG	Thymidine-phosphatidyl-guanosine
TPMT	Thiopurine methyltransferase
Tris	Hydroxymethyl aminomethane buffer
Trp	Tryptophan
TSC	The SNP Consortium
UTR	Untranslated region
VNTR	Variable number tandem repeat
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Summary

Over the past years the clinical research has rapidly advanced towards the direction of pharmacogenetics, which is defined as the study of the role of inherited variation in drug response. It appeared that individuals respond significantly different to the same drug dose; the answer lies in our genes, the differences in our genes. However, we are not talking about inconspicuous differences in the reactions to the same drug dose, the range varies between death and no effect level.

These differences are due to polymorphisms within the gene. When describing genetic polymorphisms, it is easiest to imagine two strands of DNA that differ in sequence rather than in shape.

Single nucleotide polymorphisms involve only the change of single bases in the sequence. This type of polymorphism is the most common and extremely studied form of polymorphisms in genetic research. In our studies the existence of a variable number tandem repeat plays also an important role, the variation involves individual alleles that have different numbers of repeats.

Arsenic is present in the drinking water in many parts of the world. In humans, arsenic is absorbed in the gastrointestinal tract, but it is not known as a trace element. Inorganic arsenic is methylated, and methylconjugation has been shown to be an important pathway for biotransformation of many drugs and neurotransmitters.

In this study, our goal was to analyze the genetic variations of the human arsenic methyltransferase gene with the impacts on protein function. Inorganic arsenic is methylated during biotransformation, the enzyme catalyzing this process is the arsenic methyltransferase.

We followed the model of the drug-metabolizing enzyme thiopurine methyltransferase The activity of this enzyme is known to be in human red blood cells and controlled by a common genetic polymorphism that also regulates the enzyme activity in all other tissues which have been analyzed.

The studies in which I participated involved testing the hypothesis that individual variation in the sequence or structure of the arsenic methyltransferase gene might contribute to individual differences in arsenic methylation *in vivo* and result in individual variation in arsenic toxicity and carcinogenesis. To test this hypothesis, a genotype-to-phenotype, rather than a phenotype-to-genotype strategy was used to study arsenic methyltransferase pharmacogenetics. This study initially focused on the

5'-flanking region and the single nucleotide polymorphisms in the open reading frame of this gene. However, the project started with resequencing the human arsenic methyltransferase gene. When that was done, 26 polymorphisms were observed, 3 of which were nonsynonymous coding single nucleotide polymorphisms in the open reading frame. An enzyme assay was then optimized to test the possible functions of the coding single nuclotide polymorphisms and its expressed proteins. A variable number tandem repeat was also observed in the 5'-untranslated region. A series of reporter gene constructs, incorporating portions of the 5'-flanking region as well as the 5'-untranslated region, were used to identify the promotor of the gene and to study the possible functional implications of the variable number tandem repeat.

In conclusion, we were able to prove that the differences in the methylation level are due to the polymorphisms in the arsenic methyltransferase gene. Furthermore, we showed that not only the open reading frame but also the region within the flanking region and untranslated region had an impact on the protein activity.

A major challenge for the future will be to develop ways to translate the pharmacogenetic information of the human arsenic methyltransferase gene into meaningful clinical reality and individual patient treatment.

Zusammenfassung

In den letzten Jahren entwickelte sich die klinische Forschung auch in Richtung Pharmakogenetik. Diese ist definiert als die Erforschung von unterschiedlichen Reaktionen auf dieselbe Menge eines Arzneistoffs, wofür eine genetische Ursache vorliegt. Der Grund hierfür liegt also in unseren Genen mit ihren Variationen. Eine bestimmte Dosis eines Arzneistoffs kann beim Menschen einerseits eine letale Wirkung hervorrufen oder andererseits wirkungslos sein. Diese Unterschiede werden auf Genpolymorphismen zurückgeführt. Man spricht von genetischem Polymorphismus, wenn sich Desoxyribonukleinsäure-Stränge in der Basensequenz unterscheiden. Einer der bekanntesten Polymorphismen ist die Punktmutation, welche eine Änderungen von einzelnen Basen in der Nukleotidsequenz darstellt. Diese Form wird am häufigsten beobachtet und hat in der Forschung momentan die größte Bedeutung. Als weitere Form des Polymorphismus spielt die Existenz von "variable number tandem repeats" eine wichtige Rolle. Diese Variationen schließen individuelle Allele ein, die eine unterschiedliche Anzahl von Wiederholungen einer gleichen Nukleotidsequenz haben.

In vielen Teilen der Welt findet man Arsen in erhöhten Konzentrationen im Trinkwasser. Es wird vom Menschen im Gastrointestinaltrakt resorbiert, ist aber kein Spurenelement. Anorganisches Arsen wird im Körper methyliert. Die Konjugation mit einer oder mehreren Methylgruppen spielt bei der Biotransformation von vielen Medikamenten und Neurotransmittern eine große Rolle. Die Aktivität der Thiopurin-Methyltransferase ist in menschlichen Erythrozyten messbar. Man weiß bei diesem Enzym, dass häufig auftretende genetische Polymorphismen für unterschiedliches Ansprechen auf Arzneistoffe verantwortlich sind.

Ziel der hier vorgestellten Arbeit war es, die Unterschiede im menschlichen Arsen-Methyltransferase-Gen zu analysieren und mögliche Auswirkungen auf die Enzymfunktion zu untersuchen. Wir untersuchten die Hypothese, ob individuelle Variationen des Arsen-Methyltransferase-Gens zu individuellen Enzymaktivitäten *in vivo* beitragen. Darüber hinaus wurde geprüft, ob diese Veränderungen für die Unterschiede hinsichtlich akuter Toxizität und langfristiger Tumorentstehung durch Arsen im menschlichen Körper verantwortlich sein könnten.

Wir untersuchten diese Hypothese vom Genotypen ausgehend zum Phänotypen, und nicht, wie in letzter Zeit üblich, umgekehrt. Das Projekt

konzentrierte sich am Anfang auf die 5'-flankierende Region und die Punktmutationen im offenen Leserahmen des Arsen-Methyltransferase-Gens. Zunächst wurde die entsprechende menschliche DNA sequenziert und 26 Polymorphismen waren nachweisbar: 3 Punktmutationen innerhalb des offenen Leserahmens führten jeweils zu einer Veränderung der Aminosäuresequenz. Anschließend wurde ein Testsystem entwickelt, um die Enzymaktivität von Wildtyp- und Mutantenprotein der Arsen-Methyltransferase zu quantifizieren.

Desweiteren konnte ein "variable number tandem repeat" im 5' nicht translatierten Bereich nachgewiesen. Eine Reihe von Reporter-Genkonstrukten mit unterschiedlichen Anteilen der 5'-flankierenden sowie der 5'-nicht translatierten Region wurde verwendet, um den Promotor des Gens zu identifizieren und somit mögliche funktionelle Einflüsse des "variable number tandem repeat" auf die Enzymaktivität zu untersuchen.

In der vorliegenden Arbeit konnte gezeigt werden, dass die Unterschiede im Methylierungsgrad auf die Polymorphismen im Arsen-Methyltransferase-Gen zurückzuführen sind. Nicht nur der offene Leserahmen, sondern auch die Region innerhalb des flankierenden Bereichs und des nicht translatierten Bereichs haben Auswirkungen auf die Enzymaktivität.

Die hier dargestellten Ergebnisse können in Zukunft eine wichtige Grundlage bilden, um die pharmakogenetische Information des menschlichen Arsen-Methyltransferase-Gens im Klinikalltag zu nutzen und somit eine individuelle Patientenbehandlung zu gewährleisten.

1. Introduction

Personalized medicine is a new keyword in our health system. At present time, while we consult a physician, we are all used to be asked for our body-height and weight. We know that our gender and age play an important role for the physicians and pharmacists to prescribe us the right medication and the right dosage. To go a step further, who would not be surprised if we in future were asked to show our "gene passport" when consulting a physician or pharmacist asking for medication?

1.1 Pharmacogenetics

During the last decade clinical research has rapidly advanced. Pharmacogenetics, which is defined as the study of the role of inherited variation in drug response, is the keyword. It appeared that individuals respond differently to the same drug dose; the answer for this phenomenon lies in our genes. The reactions of individuals to the same drug dose varies between death and no effect at all. And this is where the important concept of pharmacogenetics started: in the observation that there were patients with different plasma or urinary drug concentrations, followed by the realization that the biochemical traits were inherited. If we think about the consequences of these tremendous results, the future could go into the direction of individual "gene screening" to identify the right drug dose for each patient before prescribing a drug to a patient.

The idea of personalized medicine demands from the regulatory agencies to decide what genomic data they should asked for before allowing pharmaceutical companies to bring new drugs onto the market. As time advances and the prices for identifying the differences in genes fall, physicians can expect to receive more pharmacogenetic data from drug companies and moreover appropriate genetic tests for the selection of the best therapy and the right dose for the patient (Goetz et al., 2004).

The different responses of individuals regarding medical drugs are due to the variations in the genetic information lying within the deoxynucleic acid (DNA), which is the genetic material of all living organisms. It is synthesized from the basic components, the nucleotides adenosine (A), guanosine (G), cytidine (C) and

thymidine (T). The sequence of the nucleotides encodes for the various proteins and gene products.

For the average human, the process of growing from a single egg to a 60 year old adult involves 10^{17} cell divisions (Cairns, 1975). Proof-reading and repair mechanisms cannot prevent every error that might occur during that process. Therefore, mutations often occur as a result of cell replication, meiotic recombination, gene conversions, and from the effects of any number of environmental factors such as radiation or free radical damage caused by the ingestion of toxins. When a mutation occurs, a change will occur in one DNA sequence resulting in the individual having one copy of the original sequence and a second new sequence at the mutated locus. The mutations can occur either in a somatic cell or in the germ line. If they occur in the germ line, they can be passed on and effect the next generation. Every time an alternative mutation reaches a frequency of 1% or more in the population, then the locus is said to be polymorphic.

1.2 Polymorphisms

The word polymorphisms is derived from the Greek and means "having many shapes". When describing genetic polymorphisms, it is easiest to imagine two strands of DNA that differ in sequence rather than in shape. The common types of polymorphism found in the human genome can be organized into three classes:

- Repetitive elements/satellites
  Certain DNA sequences are found in multiple copies throughout the genome. Tandem repeats of a short unit are common in mammalian cells. Variability is seen when a population contains repeats of different lengths at the same genomic location.
  These sequences are called minsatellite or VNTR regions. In this case, the variation involves individual alleles that have different numbers of the repeat. The high variability in minisatellites makes them especially useful for genomic mapping.

16

●   Insertions and Deletions

Insertion and deletions are bi-allelic polymorphisms. The difference in alleles lies in the presence or absence of one or more DNA bases at the polymorphic position.

●   Substitutions

Substitutions are also most often bi-allelic polymorphisms. Alleles of this type can be recognized by a replacement of DNA bases, rather than their presence or absence.

1.2.1 Single nucleotide polymorphisms

The key defining character of single nucleotide polymorphisms (SNPs) is that alleles for these polymorphisms involve only single bases. This type of polymorphism is the most common and extremely studied form of polymorphism in genetic research. SNP alleles can exist as a result of the insertion or deletion of a single base, as well as the substitution of one base for another. In the case of substitution, the maximum number of SNP alleles is limited to four. That is because DNA is made up of only 4 different nucleotide bases: C, A, T and G.

SNPs are either created by transition (purine-purine or pyrimidine-pyrimidine) substitution or transversion (purine-pyrimidine or pyrimidine-purine) substitutions. Adenosine and guanosine are purine bases, cytosine and thymidine are refered as pyrimidine bases. The transition and transversion events appear to be more or less similar in occurrence, except for an extreme overabundance of C to T transitions. More than 70% of all SNPs found in the human genome involve C to T transitions.

This is most likely due to the chemical conversion of 5-methylcytosine residues to thymidine through a deamination mechanism (Holliday and Grigg, 1993).

The majority of polymorphic positions in the human genome are SNPs. However, the number of the SNPs released into the database seemed to increase very rapidly during the last years. As recently as April 1999, 7,000 SNPs were available in the public domain (Brookes, 1999). Less than 2 years later, in a single publication, an additional 1.42 million SNPs were released to the scientific community (Sachidanandam et al., 2001). That work was performed by the Human Genome

Project (Collins et al., 1998), and a joint academic/industrial organization "The SNP Consortium" (Masood, 1999). In 2001 the number of SNPs available in public databases was at least 3 million of the estimated 11 million SNPs thought to be present in the human genome (Kruglyak and Nickerson, 2001). This estimation would lead to a genome-wide average of 1 SNP per every 300 bp of sequence. The project of the International Human Genome Consortium was finished in 2004 and 7.2 million genetic variations and SNPs were cataloged (Thigpen, 2004).

Genetic variation is seen roughly four times less frequently in coding sequence, than in non-coding sequence (Nickerson et al., 1998; Healy et al., 1998). When a SNP is present in a coding sequence, it is also be said to be a cSNP, a coding single nucleotide polymorphism.

Many of the SNPs that occur in exons are present in the wobble position of the reading frame in a gene, these alterations in the sequence of the DNA do not change the amino acid sequence of the protein, they are synonymous or silent substitutions and have little or no effect on the gene product. Non-synonymous variants, on the other hand, cause the substitution of an amino acid on the protein level. The consequence of non-synonymous substitution on protein function varies from no effect to complete disruption of normal protein function to every gradation between those extremes.

These differences in the protein function, due to the genetic inheritance, lead to the variations of side effects that can be observed while giving same drug doses to patients.

1.3 Arsenic

1.3.1 Historical remarks

During the past century, organic arsenicals were used as drugs whereas in the Middle Ages, arsenic was commonly used as a poison to commit murder or suicide.

Arsenic containing drugs were used to heal dermatitis, asthma and epilepsy. The most famous of these compounds was arsphenanine, Ehrlichs compound 606, which was one of the first effective treatments for syphilis. But not only our ancestors saw the medical advances in the treatment with arsenic; today arsenic trioxide is used to treat leukemia. Since 2002 Trisenox$^®$ is officially accepted to treat rezidives of acute promyelocyte-leukemia. This therapy is induced when other standard treatments fail. Arsenic trioxide has effects on apoptosis and on the induction of an abnormal protein, that stops the growth and development of leucocytes (Bertsche and Schulz, 2003).

1.3.2 Effects and occurrence

Arsenic is an element with metalloid properties. It is a grayish metal, has the atomic number 33 and is found next to selenium in the periodic table, in the same group as phosphorus.

Inorganic arsenic has recently become of great public health interest, primarily because its chronic exposure has carcinogenic effects, and can cause skin- and lung cancer moreover hepatic hemangiosarcoma. Acute exposure can lead to cardiac failure, peripheral neuropathy, leukopenia and death. It has also been reported that there is a link between arsenic in drinking water and risk for the development of diabetes mellitus and hypertension (Lai et al., 1994).
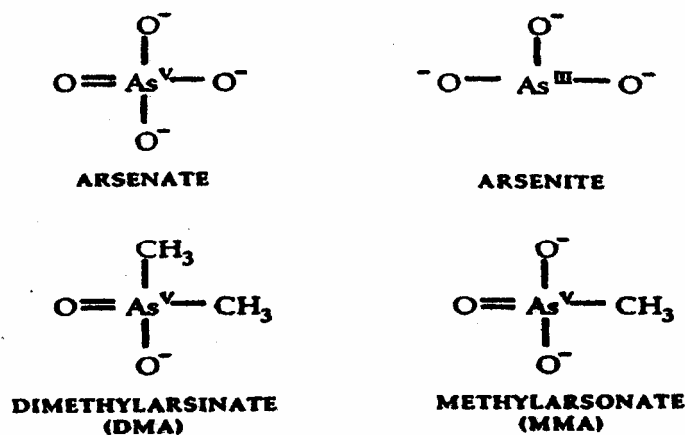
Fig. 1: Inorganic arsenate, arsenite and methylated metabolites:
monomethylarsonate and dimethylarsinate. Methylated inorganic trivalent arsenite (AsIII) is
2-10 times more toxic than methylated pentavalent arsenate (AsV).

Naturally occurring arsenic in ground water comes in two forms: arsenite and arsenate. Arsenate is toxic but arsenite is 100 times more toxic. Arsenic is present in the drinking water in many parts of the world (World Health Organization, 1981; International Agency for Research on Cancer, 1980).

Methylated arsenic compounds, monomethyl- and dimethylarsenic acid, are generated in living organisms including humans as a result of biotransformation (1.4).

Natural arsenic in ground water at concentrations over the drinking water standard of 10 µg/l is not uncommon, although in February 2002, the maximum contaminantion level was decreased from the 50 µg/l standard established in 1942 (Smith et al., 2001). Studies performed with human cell cultures have demonstrated genotoxic effects after the exposure to water containing 50 µg/l or less (National Research Council –NRC-, 2001). It was reported that an arsenic concentration of 500 µg/l in drinking water causes death from cancer in approximately 1 out of 10 subjects, while for 50 µg/l the risk is approximately 1 in 100 (NRC, 1999).

The population of Bangladesh is most likely at highest risk for the uptake of arsenic into the body. Over 30,000,000 people are counted to a potential exposed population to arsenic in the ground water with concentrations of 1-2.500 µg/l.
Although other countries like India, Vietnam and Argentina have similar ground water concentrations of arsenic, their potential exposed population is less than

6.000.000 people. In Europe, the exposed population hardly crosses 400.000 people whereas Germany does not have a potential exposed population. The ground water concentrations in Europe vary between 1 and 150 µg/l (Nordstrom, 2002).

1.4 Biotransformation

A large number of methyltransferases enzymes have already been identified and were listed in the Enzyme Nomenclature (1992). Methyltransferases cannot only catalyze the methylation of small molecules, such as arsenic, but also of macromolecules such as proteins and DNA (Aletta et al., 1998; Klein and Costa 1997; Tollervey 1996).

In humans, arsenic is absorbed in the gastrointestinal tract, but it is not known as a trace element. Inorganic arsenic is methylated, and methylconjugation has been shown to be an important pathway of the biotransformation of many drugs and neurotransmitters.

Arsenic is excreted mainly as mono-methylarsonate (MMA) and dimethylarsinate (DMA). The reduction of As^{V} to As^{III} involves the oxidation of glutathione (GSH), which is also necessary for the conversion of MMA to DMA (Thompson 1993). Urinary MMA and DMA concentrations may be used as a parameter for biological monitoring of arsenic exposure. Although methylation is typically regarded as a reaction of detoxification, some methylated compounds that contain As^{III} are more cytotoxic and genotoxic than arsenate. Methylated compounds are also better potent inhibitors of the GSH reductase (Styblo et al., 1997; Chouchane and Snow 2001). The understanding of individual variation in the enzyme, that is responsible for the methylation of arsenicals, is an important pathway for the understanding of the pathological consequences of chronic exposure to inorganic arsenic.

1.5 S-Adenosyl-L-Methionine

S-adenosyl-L-methionine (AdoMet), also well known as SAM, is the methyl donor for most methyltransferase enzymes, including the arsenic methyltransferase (AS3MT). AdoMet was discovered by Cantoni in 1953. It is synthesized *in vivo* from adenosine triphosphate (ATP) and L-methionine by methionine adenosyltransferase (MAT).

MAT catalyzes the transfer of the adenosyl group of ATP to the sulfur group of methionine. MAT also hydrolyses the triphosphate group of ATP to form pyrophosphate and phosphate (Cantoni and Durell, 1957). After the donation of the methyl group to an acceptor, AdoMet is converted to S-adenosyl-L-homocysteine.

AdoMet BIOSYNTHESIS
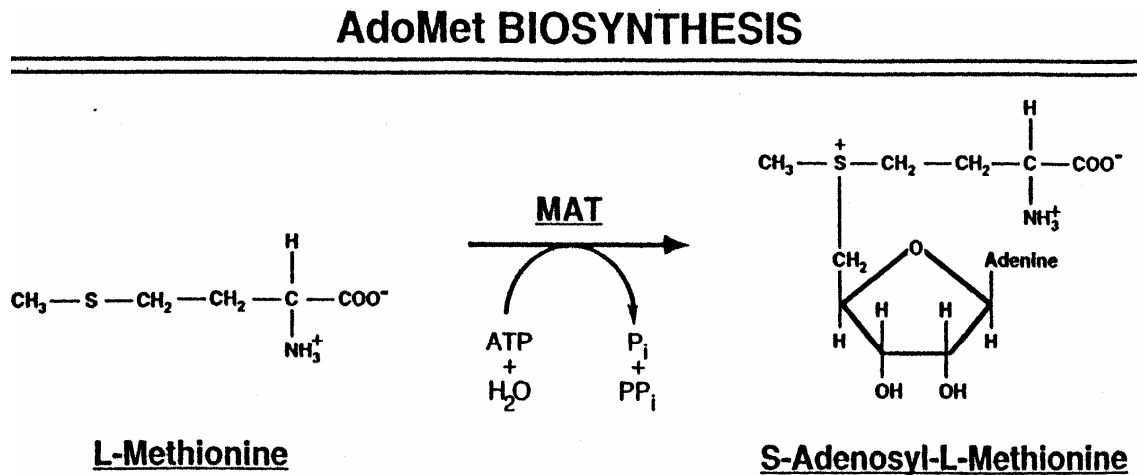
Fig. 2: Methionine adenosyltransferase (MAT) catalyzed biosynthesis of S-adenosyl-L-methionine (AdoMet).

1.6 Arsenic methyltransferase

During this study, our goal was to analyze the genetic variations of the human AS3MT gene with its impact on protein function. Briefly, inorganic arsenic, a toxin for humans, is methylated during biotransformation. The enzyme catalyzing this process is AS3MT. As we found during our study, the gene encoding this enzyme has three polymorphisms that alter the amino acid sequence. These changes in the protein structure might be responsible for the individual differences in the level of arsenic methylation.

Polymorphisms within the promotor region of the gene may influence the activity of AS3MT. Differences in the methylation of arsenic as a result of these polymorphisms in the gene might expose individuals differently to the pathological effects of arsenic in the body. The purpose of my study was to test the hypothesis that individual genetically-determined variation in arsenic methylation might contribute to individual differences in arsenic toxicity in humans. The diversity of arsenical methylation are due to the polymorphisms in the gene structure and would not only be able to influence the level of methylation of contaminated water but also the methylation of drug compounds like Trisenox[®].

AS3MT mRNA has been detected in many rat tissues. Arsenite is methylated in rat liver and kidney (Buchet and Lauwerys, 1988, Styblo et al., 1996). Arsenic methylating activity has also been reported in cytosolic fractions from mouse liver, kidney, testes and lung (Healy, 1998). Both methylated and dimethylated arsenicals have been detected in liver and kidney of arsenite treated rats (Kitchin et al., 1999). The occurrence of AS3MT mRNA in rat tissue is consistent with data indicating that the methylation of arsenic occurs in those tissues. The cloning of a rat cDNA, encoded by a gene that was originally annotated as *Cyt19,* and the identification of the protein as arsenic methyltransferase (Lin S et al., 2002), was the start for the studies of variations in this gene and individual differences in arsenic methylation.

A human liver cell line (HepG2) has also been shown to express high levels of AS3MT mRNA. These cells are efficient methylators of arsenite, producing both methylated and dimethylated products (Del Razo et al., 2001).

The *Cyt19* cDNA from rat liver cytosol was cloned by Lin et al., (2002). The human ortholog for the rat *Cyt19* gene (Genbank accession number NT 008804) is the AS3MT gene.

The AS3MT cDNA was cloned by the Mayo Pharmacogenomics laboratory of Prof. R.M. Weinshilboum, followed by an analysis of the AS3MT gene structure and sequence. Parts of these results will be presented in this work.

1.7 Thiopurine methyltransferase

Thiopurine drugs are used to treat patients with neoplasia and autoimmune disease as well as patients after transplantation (Weinshilboum, 2001).

The drug-metabolizing enzyme thiopurine methyltransferase (TPMT) emphasizes as a model the importance on pharmacogenetics of many methyltransferase enzymes. TPMT has been studied in humans from a pharmacogenetic perspective in the laboratory of Prof. R.M. Weinshilboum. TPMT catalyzes the S-methylation of thiopurine drugs such as 6-mercaptupurine (6-MP). The activity of this enzyme is known to be in red blood cells (RBC) and controlled by a common genetic polymorphism. This polymorphism regulates the enzyme activity in all other tissues which have been analyzed. Patients with inherited low levels of TPMT activity are at higher risk for thiopurine drug-induced myelotoxicity such as myelosuppression. At the same time, patients with high TPMT activity could be undertreated with the same standard dose of these drugs (Weinshilboum, 1992). Moreover, recent results indicate that TPMT could be the target for clinically significant drug interactions and that this common polymorphism might be a risk factor for therapy-dependent secondary leukemia (Weinshilboum, 2001).

After cloning and characterization of the human TPMT complementary deoxynucleic acid (cDNA), several SNPs were reported and analyzed, which seem to reduce enzyme activity. Furthermore, a polymorphic VNTR within the 5'-flanking region (FR) of the TPMT gene was found and it seems to modulate the enzyme activity (Weinshilboum, 2001).

1.8 Goals and research strategy for these studies

Our studies involved testing the hypothesis that individual variation in the sequence or structure of the AS3MT gene might contribute to individual differences in arsenic methylation *in vivo* thus, individual variation in arsenic toxicity and carcinogenesis. To test this hypothesis, we applied a genotype-to-phenotype, rather than a phenotype-to-genotype strategy to study AS3MT pharmacogenetics. The project started with resequencing the human AS3MT gene. 26 polymorphisms were observed, 3 of which were nonsynonymous cSNPs in the open reading frame (ORF). Our study initially focused on the 5'-FR and the SNPs in the ORF of the AS3MT gene.

We then started to optimize an enzyme assay to test possible functions of the coding single nucleotide polymorphisms (cSNPs) in the ORF.

A VNTR was also observed in the 5'-untranslated region (UTR). A series of reporter gene constructs, incorporating portions of the 5'-FR as well as the 5'-UTR, were used to identify the promotor of the gene and to study possible functional implications of the VNTR.

2. Materials and Methods

2.1 Materials

COS-1 cells:	American Type Culture Collection	
DMEM:	Bio Whittacker (Walkersville, MY)	
Enzymes:	Stratagene (La Jolla, CA):	PFU DNA Polymerase
	Perkin Elmer (Foster City, CA):	AmpliTaqGold DNA Polymerase
	New England Biolabs (Beverly, MA):	Restriction enzymes
	Invitrogen (Carlsbad, CA):	T4-DNA-Ligase
FBS, (PBS):	GIBCO Invitrogen (Carlsbad, CA)	
HEK/ HepG2 cells:	American Type Culture Collection	
Kits:	Qiagen (Valencia, CA):	Plasmid Maxi- and Mini Kit
	Promega (Madison, WI):	TransFastTransfection
	Promega (Madison, WI):	Dual-Luciferase Reporter assay system
	BIOsystems (Vista, CA):	101Geneclean Purification
LB-media, agar:	Bio-Whittacker (Walkersville, MY), GIBCO Invitrogen (Carlsbad, CA)	
One Shot Top 10, INVαF and DH5α:	Invitrogen (Carlsbad, CA)	
Plasmid pCR2.1:	Invitrogen (Carlsbad, CA)	
Plasmid pGL-3:	Promega (Madison, WI)	
Radioactivity:	Dupont-NEN [^{14}C-methyl]-S-adenosyl-L-methionine	
Reagents:	Invitrogen (Carlsbad, CA)	
	Roche Diagnostics (Indianapolis, IN)	
	BIOsystem 101 (Vista, CA)	
	Sigma (St.Louis, MO)	

	Baker (Phillipsburg, NJ)
	Fisher Scientific (Fair Law, NJ)
	Bio-Rad (Hercules, CA)
	Invitrogen (Carlsbad, CA)
Trypsin, α-MEM:	GIBCO Invitrogen (Carlsbad, CA)
Vector p91023:	Kindly provided by RJ Kaufmann, University of Michigan, Medical School

2.2 Equipment

Centrifuges:	Beckmann CPR, Sorvall RC-5C plus Eppendorf: Mini Spin, 5415R, 5415C
Luminometer:	TD-20/20 Tube Luminometer, Turner Designs
Polytron homogenizer:	Brinkmann instruments
Scintillation counter:	Beckmann LS-6000 SC
Shaking incubator:	Forma Scientific Model 4520
Thermal cycler:	Perkin Elmer
UV-spectrophotometer:	BioSpec- 1601 Shimadzu
Waterbath:	GCA/Precision Scientific

2.3 Growth of bacterial cultures

Bacterial cultures were grown in LB–media. Appropriate antibiotics were added for the plasmids studied: ampicillin 100 μg/ml, tetracycline 12.5 μg/ml. For plasmid-DNA preparation and isolation, the cultures were incubated for at least 16 h at 37°C and isolated as described below.

LB-media: 10 g/l trypton, 5 g/l yeast extract, 10 g/l NaCl, 15 g/l agar

2.4 DNA samples and analysis

DNA samples from African-American and Caucasian-American subjects were purchased from the Coriell Cell Repository (Camden, NJ). All subjects had provided written informed consent for the use of their DNA for research purposes and had been structured to reflect the ethnic diversity present in the United States. Sixty samples each from sample sets HD100CAU and HD100AA, collected, and anonymized from the National Institute of General Medical Sciences, were resequenced for the human AS3MT structure. The studies were reviewed and approved by the Mayo Clinic Institutional Review Board. DNA for the resequencing studies were analyzed using PolyPhred 4.0 (Nickerson et al., 1998) and Consed 8.0 (Gordon et al., 1998) programs. The human SNP databases dbSNP, HapMap and build 18 were searched to determine whether the polmorphisms that were found had been described previously. The RepeatMasker (University of Washington) program was used to screen the DNA for repeats.

2.5 Synthesis and purification of oligonucleotide primers

Oligonucleotide primers were synthesized in the Mayo Molecular Biology Core Facility with an Applied Biosystems Model 394 synthesizer. After the evaporation of ammonia by a DNA-120 Speed Vac (Therma Savant), the oligonucleotides were resuspended in H_2O.

2.5.1 Primer sequences

Table 1: Primer sequences: for the resequencing of the human AS3MT cDNA in the 5'-FR and UTR, several oligonucleotide primers were designed in the Mayo Molecular Biology Core Facility. Cleavage sites were generated for future digestions by restriction enzymes.

Location	Primer	Tm	bp	Sequence	Cleavage site
Exon 11	AMTF R1156	63	28	CTACTGTGGTCCCCTGGTTGGCTATAGA	--
Exon 1	Arsenic F(-78)	62	23	GAGACAGTGAGTGCGCGCCCTGA	--
5'-Flanking reg.	hAMTU F-312AC	64	37	AAGAAGggtaccACGAGATTTATCCGTGAA AATCGCA	*Acc65I*
Intron 1	hAMTI1 R32XHO	70	31	AAGAAGctcgagAGGGAAGGGGCTGGGGGCT	*Xho1*
5'-Flanking reg.	R91	71	33	AAGAAGctcgagTCATGATGACGGCGGGCTC CCGCA	*Xho1*
5'-Flanking reg.	UFF-113AMT	73	35	AAGAAGggtaccGCTGCGGGAGCCCGCCGTC CTGA	*Acc65I*
5'-Flanking reg.	VNTR F(-90)	71	37	AAGAAGggtaccGTCGCAGGCCGAGGAGACA GTGAGT	*Acc65I*
Intron 1	VNTR R(16)	71	36	AAGAAGctcgagCAGGGCGCGCACTCACTGT CTCCT	*Xho1*
pGL-3	rev402	64	24	CGCGGGCGCAACTGCAACTCCGAT	--

Calculation of Tm (°C) = (G/C) x 4°C + (A/T) x 2°C – mismatching bp x 4°C

2.6 Quantification of nucleic acids

Nucleic acid absorptions were determined with a UV-spectrophotometer. For these measurements, the extinction of the DNA solution was quantified at 260 and 280 nm (1 cm cuvette). Absorption of 1 equals 50 µg/ml double stranded DNA, 33 µg/ml single stranded DNA or 40 µg/ml RNA. The ratio of the extinction at 260 and 280 nm for a pure DNA-solution is 2.

2.7 Preparation of plasmid-DNA

2.7.1 Maxi-purification

Three different cell lines were used for transfection: COS-1 cells, an african green monkey kidney cell line; HepG2 cells; and HEK293 cells, a human embryotic kidney cell line. As a first step, plasmid DNA was purified from a 500 ml bacterial culture based on the protocol for the Qiagen Plasmid Maxi Kit:

Bacterial were harvested upon centrifugation (15 min, 6000 x g at 4°C) and the remaining bacterial pellet was resuspended in 10 ml of buffer P1. Cells were lysed in 4 ml of buffer P2. The lysate was neutralized by the addition of 10 ml of buffer P3, and incubated on ice for 20 min. After centrifugation (30 min, 20 000 x g, 4°C) a clear plasmid DNA lysate was produced for loading onto a Qiagen-tip by gravity flow. To equilibrate the Qiagen tip, 10 ml of buffer QBT was applied to the tip. The plasmid DNA bound to the tip, while degraded RNA and cellular proteins were not retained and appeared in the flow-through fraction. The Qiagen tip was washed with 2x30 ml of buffer QC. To elute the DNA, 15 ml of buffer QF was applied.

To precipitate DNA, 10.5 ml isopropanol was added to the eluted DNA and the sample was centrifuged (30 min at 15 000 x g, 4°C). The DNA pellet was washed with 5 ml of 70% ethanol at room temperature and centrifuged again (15 000 x g, 10 min). After air-drying for 10 min, the DNA was redissolved in 1.5 ml of TE buffer.

Reagents:

Buffer P1:	50 mM Tris-HCl pH 8, 10 mM EDTA, 100 µg/ml RNase A
Solution P2:	200mM NaOH, 1% SDS
Solution P3:	3.0 M $KC_2H_3O_2$ pH 5.5
Buffer QBT:	750 mM NaCl, 50 mM MOPS pH 7, 15% $CH_3CHOHCH_3$, 0.15% Triton X-100
Buffer QC:	1.0 M NaCl, 50 mM MOPS pH 7, 15% $CH_3CHOHCH_3$
Buffer QF:	1.25 M NaCl, 50 mM Tris-HCl pH 8.5, 15% $CH_3CHOHCH_3$
Buffer TE:	10 mM Tris-HCl pH 8, 1 mM EDTA

2.7.2 Mini-purification

The Mini-purification protocol involved a purification volume of 1.5 ml of bacterial culture. The culture was centrifuged (2 min at 10 000 x g); the supernatant discarded

and the pellet was resuspended in 200 µl Cell Resuspension solution, buffer P1. After vortexing, 200 µl Cell Lysis solution, buffer P2, and Neutralization solution, buffer P3, were added, and the lysate was centrifuged (5 min at 10 000 x g).

For plasmid DNA purification, 1 ml resin was added to each minicolumn assembly, followed by the cleared lysate. To wash the minicolumns, 2 ml column wash solution was added; and the sample was centrifuged (2 min, 10 000 x *g*). To elute DNA, 50 µl nuclease-free water was applied to the minicolumn, incubated for 2 min at room temperature and centrifuged (20 sec, 10 000 x g).

Reagent compositions: 2.7.1
Wash solution: 80 mM $KC_2H_3O_2$, 8.3 mM Tris-HCl pH 7.5, 40 µM EDTA

2.8 Modification of DNA: Adding an "A" overhang to the end of a DNA fragment

For the addition of A's, the DNA nucleotide base adenosine, DNA 5'ends fragments, 2 µl AmpliTaqGold buffer, 2 µl (200µM) deoxyribonucleotide triphosphate (dNTP), 0.2 µl AmpliTaqGold enzyme and 2 µl H_2O were added to each polymerase chain reaction (PCR) product. The mixture was then incubated in a thermal cycler at 94°C for 12 min, followed by 72°C for 10 min.

2.9 DNA Preparations

2.9.1 Digestion

DNA from plasmids was incubated with the restriction enzymes *XhoI* (20 000 units/ml) and *Acc65I* (10 000 units/ml) for 4 h at 37°C. 13 µl H_2O, 0.5 µl of each restriction enzyme, 1 µl enzyme-buffer and 5 µl DNA was combined in a total volume of 20 µl. For further analysis, probes were visualized and purified from agarose gels.

2.9.2 DNA-sequencing

Amplicons were sequenced with an ABI377 DNA sequencer (Applied Biosystems, Foster City, CA) in the Mayo Molecular Biology Sequencing Core Facility using Big Dye (Perkin Elmer) dye primer chemistry. All samples were sequenced on both strands.

2.9.3 Electrophoresis

For the separation of DNA fragments of different sizes, variable concentrations of agarose gels were made varying between 0.5% (separation of fragments > 4000 bp) and 5% (separation of fragments < 100 bp). 10% TAE (Tris-acetate-EDTA) buffer pH 8 and 1% high melting agarose were heated and dried in a gel chamber. The gels were loaded with 2 µl DNA sample buffer and 11 µl DNA for each sample. The marker contained 2 µl dye, 8 µl H_2O and 3.5 µl 1 kb ladder. Electrophoresis was started immediately at 20 W and usually ran for 1.5 h. To make the different DNA fragments visible, the gel was stained with ethidium bromide (10 mg/ml) for 10-20 min. The intensity and lengths of the bands were visualized by UV-light.

Buffer TAE: (2L) 96.8 g Tris, 21.8 g acetic acid, 5.84 g EDTA
DNA sample buffer: 0.5% bromphenol blue, 0.5% xylencyanol, 0.5% glycerol

2.10 Isolation of DNA fragments

DNA fragments were isolated based on the protocol with the BIO 101 Geneclean Kit: bands were excised from the agarose gel and applied into a small tube; three volumes of 6 M NaI were added and samples were incubated at 55°C to melt the gel. After applying glassmilk suspension (aqueous suspension of proprietary silica matrix) the glassmilk/DNA complex was centrifuged (9000 x g, 15 sec). To wash the pellet, 300 µl of Newwash solution was applied 3 times to the tube. The sample was centrifuged (9000 x g), and the liquid was discarded. For elution, purified DNA was resuspended in 20 µl of H_2O.

Newwash solution: 100 mM NaCl, 50% CH_3CH_2OH

2.11 Cloning of DNA fragments

2.11.1 Overview of vectors and expression systems

Competent *E.coli* strains INVαF, DH5α and Top10 were used to perform transformation. DNA sequences to be studied were cloned into the plasmid pCR2.1, followed by subcloning into the plasmid pGL-3. To create eukaryotic expression constructs, sequences were subcloned into the expression vector p91023(b) (Kaufman, 1985) and inserts were verified by sequencing on both strands. Expression and reporter gene constructs were expressed in HepG2, HEK293 and COS-1 cells.

2.11.2 Ligation and transformation

For ligation, the vector and fragments were combined at a concentration ratio of 1:3 (vector:insert) in a final ligation volume of 10 µl. The mixture was incubated overnight at 14°C in a thermal cycler with 1 µl of T4-DNA-Ligase as well as incubation buffer supplied by the company.

Competent E.coli cells were thawn on ice and 2 µl of the belonging ligation reaction mixture was added to each vial, followed by incubation on ice for 30 min. The mixture was heat shocked at 42°C for 30 sec, stored on ice and 250 µl SOC

media was applied at room temperature. Vials were then shaken horizontally at 37°C for 1 h at 225 rpm in a shaking incubator. Cells were plated on agar plates (LB media with 15 g/l agar) with corresponding antibiotics.

SOC media: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM glucose

2.11.3 Transfection

24 h before the transfection of COS-1-, HEK-, and HepG2 cells, a vial of Promega Transfast reagent was thawn at room temperature. 400 µl of nuclease-free water was added, vortexed and the reagent was stored overnight at –20°C.

For transfection, the steps described in the Promega TransFast Transfection protocol, were followed. Briefly, as a first step the DNA concentration was determined spectophotometrically. 10 µg of plasmid DNA was added to 30 µl of Transfast reagent and 5 ml of serum free Dulbeccos's Modified Eagel medium (DMEM). After the transfection time of 1 h, 12 ml of DMEM was added to the transfection mixture and incubated for 48 h.

2.11.4 Cell harvesting

2.11.4.1 Cell harvesting for enzyme activity determination

Cells were cultured for 48 h at 37°C. For harvesting the cells, the media was sucked off, 0.25% trypsin was added to the plates; fresh media was applied to each plate; and the mixture was centrifuged at room temperature (5 min at 1000 x g). To resuspend the cells, 10 ml of phosphate buffered saline (PBS), was added and the sample was centrifuged for 5 min at 1000 x g. After removing PBS, 6 ml homogenization buffer was applied. Before centrifugation for 15 min at 15 000 x g, the cells were homogenized with a Polytron homogenizer (2 x 15 sec). After centrifugation at 100 000 x g for 1 h at 4°C the high speed supernatant (HSS), cytosol preparations, were recovered and stored frozen in aliquots at –80°C.

PBS: 137 mM NaCl, 2.6 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7
Homogenization buffer: 10 mM Tris-HCl, pH 7.6, 250 mM sucrose, 0.5 mM reduced GSH, 0.2 mM PMSF

2.11.4.2 Cell harvesting for the luciferase assay

After removing the growth medium from cultured cells, PBS was added to wash the surface of the culture vessel followed by aspiration. The recommended volume of 1x Passive Lysis buffer (Promega) was added to each culture well. For a 6-well culture plate, 500 μl was recommended. The culture plates were then placed on a rocking platform at room temperature for 15 min, before following the steps for the luciferase assay (2.15.2).

2.11.5 Subcloning of the insert from vector pCR2.1 into the vector pGL-3

Both vectors were digested as described above with appropriate restriction enzymes (*XhoI* and *Acc65I*). pCR2.1 (3.9 kb) was digested to remove the insert (between 325 and 397 bp), and pGL-3 was digested to open the vector at the specific restriction sites. The band of the vector pGl-3 (4.9 kb) and the insert from pCR2.1 were excised

from the gel, each insert of 300 bp was combined with pGl-3 and ligated as described above.

2.11.6 Screening the pGL-3 vector for the insert

Colonies were picked from plates with a sterilized toothpick and rinsed with H_2O into a PCR plate. Colonies in the PCR plate were incubated for 5 min at 99°C. PCR was set up and performed with primers that would amplify the insert in the vector. To determine whether subcloning into the vector had been successful, DNA fragments were run on an agarose gel as described above.

2.12 Preparation of LB-agar with ampicillin and X-gal

LB-agar was autoclaved for 20 min at 121°C and incubated in a waterbath for 20 min at 55°C. 1 µl ampicillin solution (100 µl/ml) per 1 ml LB-agar was added for the selection of bacterial colonies. 25 ml of LB-agar was placed in each culture plate and air dried.

Before transferring the insert ligated in pCR 2.1 to the plates, 40 µl of X-gal (1 mg/ml) was pipetted onto each plate. The grown cultures were plated on agar and selected using the blue-white screen. The insertion of DNA into the pCR 2.1 vector interrupts expression of the β-galactosidase gene. The white color of the colonies indicates that an insert interrupted the gene for β-galactosidase and X-gal cannot function as a substrate. On the other hand, a failure of insertion results in blue colonies. The pGL-3 vector does not have the β-galactosidase gene. Since ampicillin was applied to the agar, only bacteria containing the vector and, therefore, the ampicillin resistance gene would be able to grow on the agar.

2.13 PCR amplification

PCR was performed with a thermal cycler. For a 50 µl reaction, the following ingredients were added:

DNA (10 pg-1 µg)	1 µl
Cloned / Native PFU, AmpliTaq Gold Buffer	5 µl
dNTPs (10mM)	2.5 µl
H_2O	40 µl
Primer each (1 µg/ml)	1.5 µl
$MgCl_2$/DMSO	0.5 µl
PFU, AmpliTag Gold Polymerase	0.8 µl

The number of cycles varied between 25 and 30.

Reaction Conditions:

Predenaturation:	96°C	120 sec (PFU), 12 min (AmpliTaqGold)
Denaturation:	96°C	30 sec
Annealing:	(Tm of the oligonucleotides -2°C)	30 sec
Elongation:	72°C	30 to 60 sec
Extension:	72°C	5 to 10 min
Stop:	4°C	

2.14 Cell culture

Tissue culture:

COS-1 cells, HepG2 cells and HEK293 cells were cultured at 37°C with 5% CO_2 in an incubator. DMEM media for COS-1 and HEK cells and Modified Eagel medium (α-MEM) for HepG2 cells was combined with 50 ml of fetal bovine serum (FBS). The cells were seeded every 3-4 d. To remove cells from plates and flasks, 0.25% trypsin was added to the cells and incubated at 37°C for 5-10 min. DMEM for COS-1 cells and MEM for HepG2 cells, both media included FBS, was then applied to the trypsinized cells and as a last step, DMEM (MEM) with FBS was added to a new plate or flask and the trypsinized cells were transferred to tissue culture plates. The cells were not split more than 20 times.

2.15 Assays

2.15.1 Arsenic methyltransferase assay

The arsenic methyltransferase assay was performed based on the optimization as described by Zakharyan et al., 1995. Separating, identifying, and measuring the products of the reaction using chloroform extraction validated the assay. The enzyme activity was measured in human liver cytosol and COS-1 cytosol, after transfection with constructs that contained the AS3MT WT sequence as well as the three cSNPs at nucleotide positions 917, 860 and 517.

Incubation conditions: 30 μl 1 M Tris-HCl pH 8, 10 μl 0.1 M GSH, 5 μl 0.05 M MgCl$_2$, 5 μl ^{14}C labeled 500 μM AdoMet (the methyl donor with a specific activity of 2.4 x 10^8 Ci/mmol) and enzyme preparations were combined in a final volume of 250 μl. For each sample, five tubes were used, two "blanks" with 2 μl H$_2$O instead of substrate and three "actives" with 2 μl 12.5 mM NaAsO$_2$ as a methyl acceptor. The reaction was started by adding a dilution of 100 μl of enzyme HSS from the specific sample studied and 98 μl homogenization buffer (2.11.4.1). Samples were then incubated for 60 min in a 37°C shaking water bath, and the reaction was stopped by adding 750 μl 12 M HCl.

Extraction procedure: for the extraction of methylated products, 750 μl CHCl$_3$, 10 μl KI (40%) and 20 μl K$_2$Cr$_2$O$_7$ (1.5%) were added. The contents were vortexed for 3 min, followed by centrifugation for 3 min at 1500 x *g*. The acidic aqueous upper phase containing AdoMet was removed and the lower chloroform layer was washed twice with a mixture of 250 μl H$_2$O, 5 μl 40% KI and 750 μl 12 M HCl. After each wash, the tubes were vortexed and centrifuged for 3 min at 1500 x *g* and the upper aqueous layer was removed. The methylated arsenic compounds, which remained in the chloroform layer, were back-extracted by adding 1 ml H$_2$O, followed by vortexing for 3 min and centrifugation for 5 min at 1500 x *g*. To measure the activity of methylated arsenic compounds, 0.75 ml of the upper aqueous layer was transferred to a scintillation counting vial, and 5 ml liquid scintillation cocktail (BIO- Safe 2) was added. Counts per minute (cpm) were determined in a liquid scintillation counter.

2.15.2 Luciferase assay

The luciferase assay was performed as described in the Dual-Luciferase Reporter assay system protocol (Promega). Briefly, 100 µl LAR II (luciferase assay reagent) containing Firefly luciferase was added to a tube that contained 20 µl of the cell lysate. After the initial measurement of Firefly luciferase activity with a luminometer, 100 µl Stop&Glo reagent was added (Turner Designs), and the second luminometer measurement of Renilla luciferase activity was performed. This latter value was divided by the Firefly activity to correct for transfection efficiency. Results are reported as the ratio of Firefly luciferase light units to light units for Renilla luciferase.

3. Results

3.1 Human AS3MT gene structure

The human AS3MT gene is located on chromosome 10q24. The cDNA is 1602 bp in length, with an ORF of 1125 bp. The resulting protein of the AS3MT gene has 375 amino acids. The AS3MT gene consists of 11 exons, 10 introns and is approximately 32 kb in length. The first exon, 78 bp in length, encodes mainly the 5'-UTR and the final nucleotide of that exon, adenosine, is the first nucleotide of the translation initiation codon. Exon 11, 528 bp in length, mainly encodes the 3'-UTR.

The nucleotides of the ORF of AS3MT are 81% identical and the amino acid sequences are 76% identical with those for the Cyt19 gene, the rat ortholog for the human AS3MT gene.
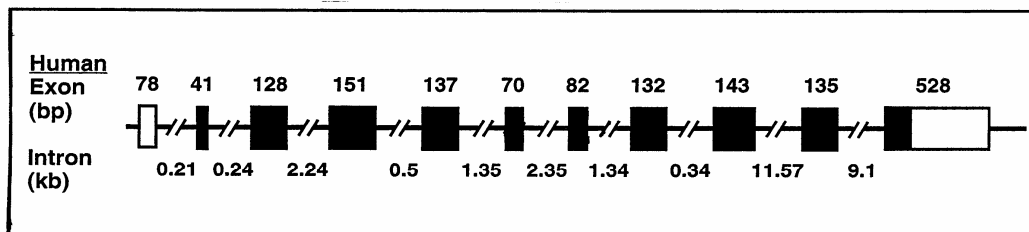
Fig. 3: Human AS3MT gene structure. Exons are represented as rectangles, with black rectangles indicating coding exons and white rectangles showing UTR sequences.

During gene resequencing 26 polymorphisms were found, 3 of the coding region SNPs were non-synonymous. These SNPs alter the sequence of amino acids of the protein expressed by the AS3MT gene.

The 3 non-synonymous cSNPs found in the ORF in the AS3MT gene changed the amino acid at the following positions:

- Exon 6, nucleotide <u>517</u>:   a change from C to T, resulting in an amino acid exchange from Arg173 to Trp173 (Variant 517);

- Exon 9, nucleotide <u>860</u>:   a change from T to C, resulting in an amino acid change from Met287 to Thr287 (Variant 860);

40

- Exon 10, nucleotide <u>917</u>:   a change from C to T, resulting in an amino acid exchange from Thr306 to Ile306 (Variant 917).

The cSNP at nucleotide position 860 in exon 9 showed the highest frequency, with an allele frequency of 10% in both Caucasian-American and African-American subjects, cSNP at nucleotide 517 was only found in African-Americans with a frequency of less than 1%, and cSNP at nucleotide 917 was observed only in Caucasian-Americans with a frequency that was also less than 1%.

A VNTR with different repeats of a nucleotide sequence, varying between 2, 3, and 4 repeats of a 36- and 35-nucleotide sequence, is located in the portion of the gene encoding the FR and the 5'-UTR. The UTR region is transcribed, but not translated.

The repeat consists of 1, 2, or 3 copies of the 36 bp repeat, the last copy of the repeat has 35 bp. 120 samples were obtained from the Coriell Cell Repository, 60 African-American and 60 Caucasian-American samples were analyzed for the VNTR length. Allele combinations for the VNTR number are 2/2, 2/3, 2/4, 3/3, with only one out of 120 analyzed samples having the genotype 3/4. The 4 repeat allele was only observed in African-American subjects. The two repeat VNTR is linked to a SNP located 23 nucleotides upstream from the beginning of the VNTR, equivalent to (-114)bp from the ATG start codon.

The AS3MT resequencing data have been deposited in the NIH database PharmGKB and has the accession ID PA128747780.

3.1.1 NCBI AS3MT gene annotation

The National Center for Biotechnology Information (NCBI) annotation for the AS3MT gene (NM_020682) is based on an exon prediction software to the early version of chromosome 10. The ORF nucleotide 1057 in the human resequenced AS3MT is missing in the NCBI annotation. It results in a frameshift and truncation of the protein within exon 10, exon 11 converts to 3'-UTR. Moreover, there are 4

variant nucleotides in exon 5 in the NCBI annotation. The following alterations result from these changes (Li J et al., 2005):

- Ile132Phe;
- Tyr135Asn;
- Gly140Ala.

After resequencing the human AS3MT gene we could not observe those polymorphic nucleotides in our laboratory. Furthermore, deletion of nucleotide 1057 was not present in any of the 120 DNA samples studied.

3.2 Methylation activity

3.2.1 Detection of methylation activity

We first focused on the cSNPs in the ORF of the human AS3MT gene.

Prior to this assay, AS3MT and mutated types were cloned in to the eukaryotic expression vector p91023(b):

- WT sequence;
- Nucleotide <u>517</u> SNP C/T in exon 6;
- Nucleotide <u>860</u> SNP T/C in exon 9;
- Nucleotide <u>917</u> SNP C/T in exon 10.

The assay for the detection of methylation activity of arsenic compounds was first published by Zakharyan et al., 1995. This group developed an assay, which measured the activity of a native arsenic methylation enzyme, purified from rabbit liver. Prof. R.M. Weinshilboum and his lab characterized the human cDNA for the AS3MT gene, the ortholog for the rat AS3MT (Cyt19) gene, described by Li et al. 2002.

We now decided to optimize the conditions of the methylation assay first and then to study the contribution of variation in this gene to individual differences in arsenic methylation.

The four constructs in the eukaryotic expression vector p91023(b) were transfected into COS-1 cells (2.11.3), which is a mammalian cell line and could ensure that appropriate mammalian post-translational modifications and protein degradation systems would be present. For the assay of the methylation activity, organic solvent extraction procedure as described in the Methods section (2.15.1) was used to separate radioactive AdoMet from radioactive MMA and DMA (Fig. 1), the compounds produced by AS3MT.

After several expression of the AS3MT WT protein and 3 variant allozymes, it appeared that the methylation activity of the cSNP at position 860 expressed a protein with the highest methylation value, although the activity levels varied among transfections. It was obvious that this common cSNP in exon 9 resulted in the expression of a protein with highest methylation activity compared to the expression of AS3MT WT or the other two cSNP's. The variant 517 protein, showed little and variant 917 almost no methylation activity.

The signal, the average of the blanks subtracted from the average for the actives, for the AS3MT WT protein varied between 40 and 60 cpm, while the signal for the AS3MT variant with the highest frequency (10% in Caucasian- and African-Americans) variant 860 was between 180 and 250 cpm. The other two expressed proteins, variant 517 and variant 917, did not show a significant activity for the methylation of arsenic compounds, with a ratio, the signal divided by the average of the blanks, hardly greater than three.
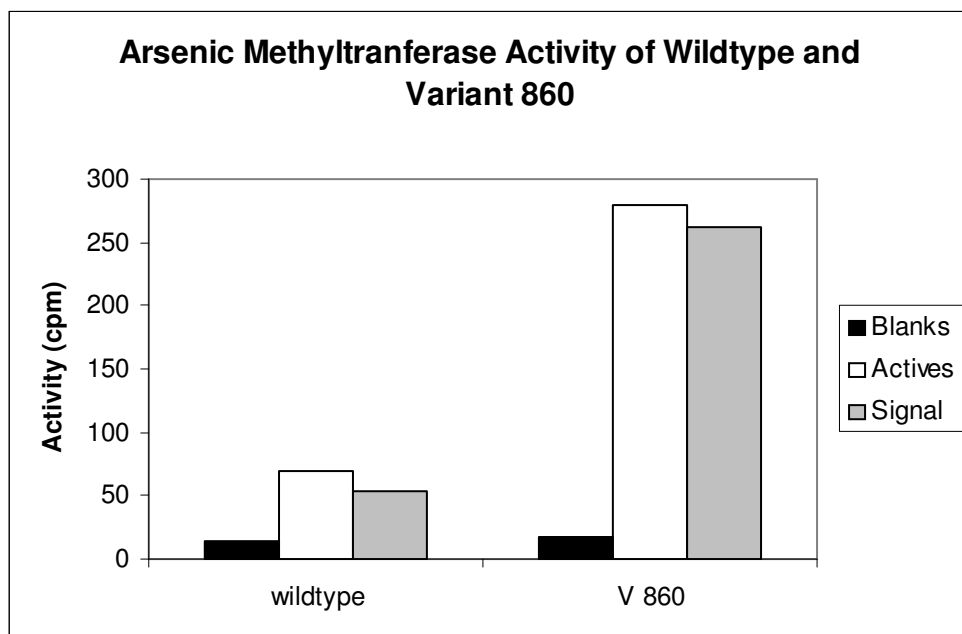
Fig. 4: AS3MT methylation activity for the expression of recombinant WT and variant 860 (V 860). "Blanks" contained 2 μl H$_2$O instead of substrate, the "actives" had 2 μl 12.5 mM NaAsO$_2$ as methyl acceptor.

For the correction of possible AS3MT enzyme activity in COS-1 cells, an "empty" p91023(b) vector was transfected into COS-1 cells.

The cpm values in the HSS from COS-1 cells, as a negative control, did not show significant arsenic methylating activity.

3.2.2 Optimization of the assay

For the optimization of the assay for the methylation of arsenic compounds, several conditions were tested. We started to optimize the assay with the AS3MT variant 860, the expression of the cSNP in exon 9, because this protein showed higher methylation activity than the AS3MT WT and the other variants.

The time course, optimized with the variant 860, and tested for 15, 30, 45, 60 and 75 minutes, had the highest ratios -between 14 and 16-, with an incubation time between 45 and 75 minutes for the reaction in a shaking water bath at 37°C.

The enzyme dilution curve, tested for 0, 20, 40, 60, 80, 100 μl HSS of recombinant enzyme of the AS3MT variant 860, was tested with a concentration range of 5-50 μg. The highest methylating activity was found with 100 μl of HSS.

Fig. 5: Protein dependence curve for the AS3MT variant 860. Recombinant AS3MT enzymes were assayed from 5-50 μg.
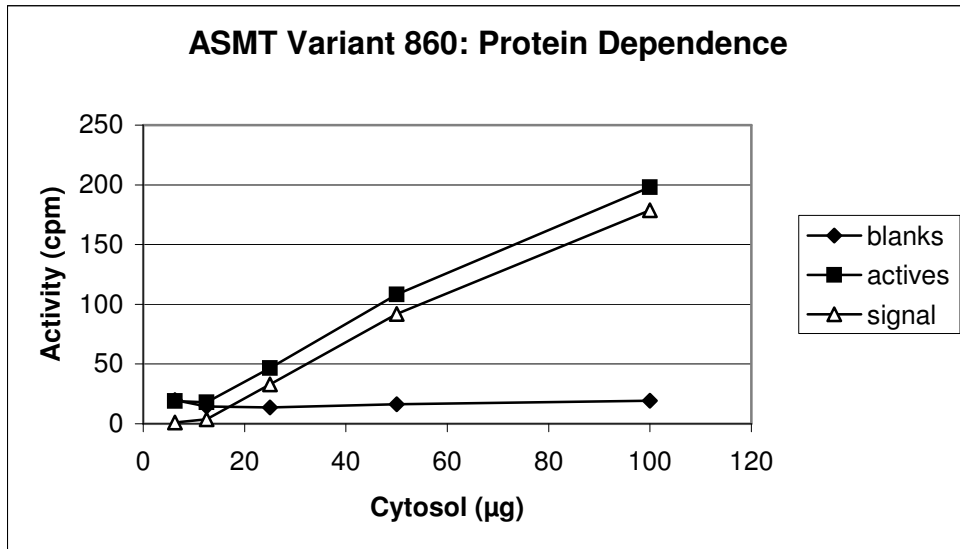
For the substrate curve, concentrations of 0.008, 0.08, 0.8, 8, 80 and 800 μM arsenite were assayed.
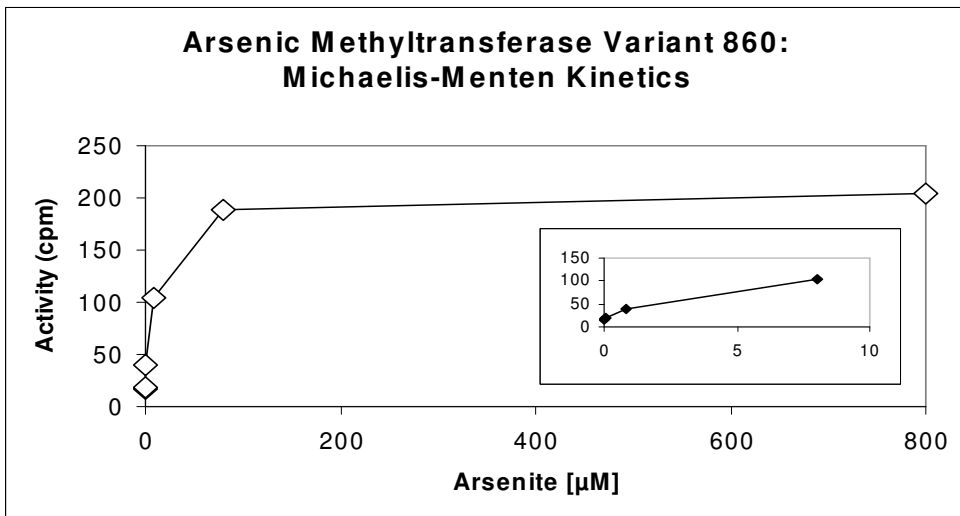

Fig. 6: Substrate curve for the AS3MT variant 860.

The cpm's for the different values of the arsenite concentration are shown underneath.

Table 2: Arsenite concentration in μM and activity in cpm.

Arsenite [μM]	Cpm	Arsenite [μM]	Cpm
0.0008	0	8	88
0.08	2	80	171
0.8	23	800	188

The highest activity was almost reached with an arsenite concentration as a substrate of 80 μM. At this point we were working with a concentration of 12.5 mM of sodium arsenite, but an activity level was already shown at lower substrate concentrations.

The maximum velocity of the enzyme is $V_{max,}$ here 188 cpm. The K_m value is roughly an inverse measure of the affinity or strength of binding between the enzyme and its substrate. This means in our case: between the AS3MT variant 860 and the substrate arsenite. The K_m resulted here in about 7μM, which means that the affinity between the substrate and enzyme is high.

3.2.3 Controls and transfection efficiency

In an attempt to find a positive control for the assay, beagle liver and kidney HSS were tested for methylation activity. HSS was pooled from 10 samples. The kidney pool showed almost no activity, and the liver pool had a signal-to-noise ratio of only two.

A transfection corrected with β-Gal for the WT and the three variants, showed lower counts for the methylation of arsenic compounds. It seemed that the cotransfection might have inhibited the enzyme activity. The ratio for the variant 860, for example, dropped from 15 down to 3.

3.2.4 Outcome and further steps

Due to much variability in the results, a failure to find a positive control for the assay, and neither stable reproducibility nor high methylation signal, moreover a WT protein with low cpm, a future goal would become to determine whether the assay is not sensitive enough to detect the arsenic methylation or whether methodical factors influenced our findings.

At that time the Weinshilboum lab was not able to determine which changes would lead us to reproducible results or which technical equipments we should change to get stable methylated activity of the four AS3MT gene constructs. In April 2003 we decided to take those as preliminary studies and to order an antibody directed against AS3MT amino acids 341-360 for western analysis. My part became to focus on the VNTR region and to test the influences on the polymorphisms in the promotor region on the gene activity.

When I left the states in September 2003 (my appointment at the Mayo Clinic and my visa expired), Tom Wood was still working on the assay, the method remained the same, but ^{14}C labeled SAM was exchanged against ^{3}H labeled SAM. Moreover, some concentrations of the reagents were changed: 0.10 M Tris-HCl buffer instead of 1 M, 4 mM glutathione instead of 0.1 M and 1 mM $MgCl_2$ instead of 0.05 M.

In march 2006 the results were published in the Journal of Biological Chemistry (Wood et al., 2006). A difference to my data and the data presented was, that this paper included the new results from the haplotype analysis (Schaid et al., 2002), which had shown that the cSNP at nucleotide 917 (Ile306) was only observed in the presence of the cSNP at nucleotide 860 (Thr287).

The results for the expression of the WT and four variant proteins are presented in Fig. 7. The results of the arsenic methyltransferase activity showed the same pattern as the preliminary studies.
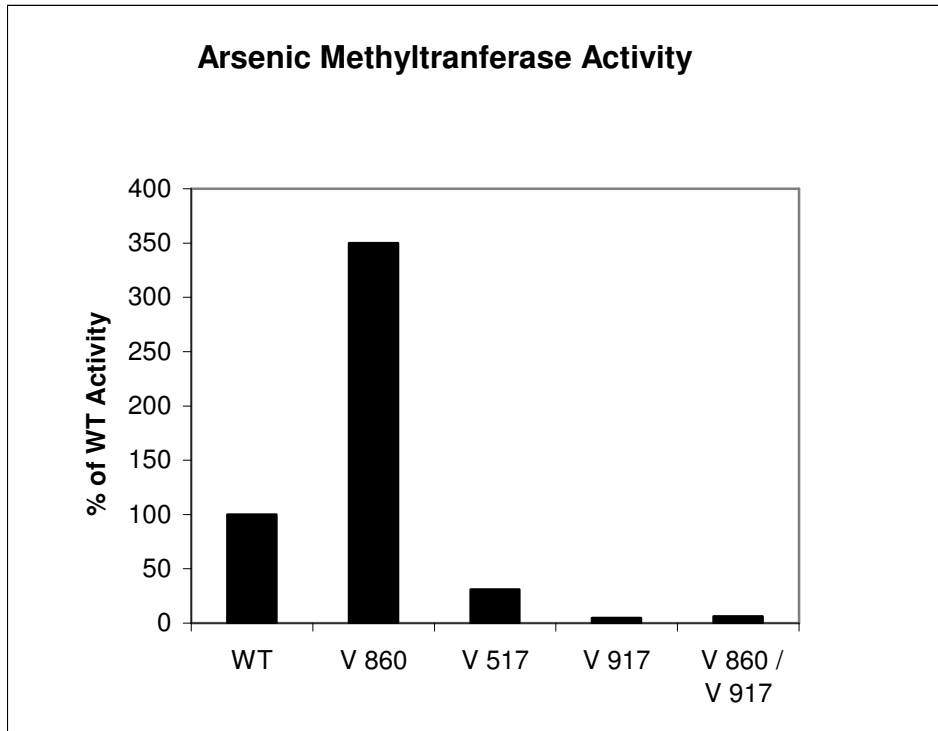
Fig. 7: Final results of the AS3MT activity as percentage of the AS3MT WT activity with variant 517 Trp173, variant 860 Thr287 and variant 917 Ile306. The activity of the WT enzyme was taken as 100%.

After the correction for transfection efficiency and six transfections the following methylation activities of the AS3MT WT and variant proteins with sodium arsenite as a substrate were shown:

- the variant 517   Arg173 to Trp173 $31\pm2.6\%$;
- the variant 860   Met287 to Thr287 $350\pm89\%$;
- the variant 917   Thr306 to Ile306 $3.2\pm2.1\%$;
- the variant 860 linked to the variant 917 $6.2\pm1.9\%$.

Since the proteins of the WT and 3 variants of AS3MT, showed different methylation activity levels, western analysis for the arsenic methyltransferase was done to measure the amount of AS3MT protein. A rabbit polyclonal antibody used to detect AS3MT was directed against amino acids 341-360, it did not include any of the amino acids which are altered in the variants. The corrected results for transfection efficiency showed the following levels of the WT immunoreative protein:

- the variant 517   Arg173 to Trp173 $20\pm0,5\%$;

- the variant <u>860</u>   Met287 to Thr287   190$\pm$14%;
- the variant <u>917</u>   Thr306 to Ile306   4.4$\pm$0.6%;
- the variant <u>860</u> linked to the variant <u>917</u>   7.9$\pm$3.1%.

3.3 VNTR studies

The VNTR is located in the portion of the gene encoding the 5'-UTR, the UTR is transcribed but not translated. The VNTR is a repeat, which consists of 36 bp and 35 bp repetitive elements. It consists of 1, 2, and 3 copies of the 36 bp sequence, and one repeat consists of 35 bp. We drew our attention to the following ideas:

- analyzing whether the VNTR alleles fit the Hardy-Weinberg hypothesis or not;

- to examine 120 Coriell samples concerning the SNP in the FR, the number of repeats in the VNTR and the cSNP at position 860 in exon 9;

- to design several reporter gene constructs to study the possible functional influences of the VNTR.

3.3.1 Hardy-Weinberg equilibrium analysis for the VNTR alleles

Evolution involves changes in the gene pool. A population with Hardy-Weinberg equilibrium shows no change, and genotype frequencies are constant from generation to generation. Populations are able to maintain a reservoir of variability so that, if future conditions require it, the gene pool can change. If recessive alleles would always tend to disappear, the population would soon become homozygous. Under Hardy-Weinberg conditions, genes that have no present selective value will nonetheless be retained. We set out to test whether the VNTR alleles fit the Hardy-Weinberg equilibrium or not.

Caucasian-American subjects (N = 60):

Alleles (repeats):	2/2	2/3	3/3
Observed numbers:	11	23	26
Expected numbers:	8.4	28.1	23.4

Chi-square = 1.99

African-American subjects (N = 60):

Alleles (repeats):	2/2	2/3	2/4	3/3	3/4	4/4
Observed numbers:	6	26	7	20	1	0
Expected numbers:	8.4	25.1	3	18.7	4.4	0.2

Chi-square = 9.11

The Chi-square value was calculated with the Georgetown University Web Chi-Square Calculator. This statistical test that can be used to determine whether observed frequencies are significantly different from expected frequencies. Chi-square is larger if the seize of the deviation or the number of groups is larger. It is smaller when the expected number is larger. In general, if the Chi-value is larger, the deviations are more likely to be significant, and the data is less likely to fit the expectations.

The value 3.84 equals the 5% limits for maintaining the hypothesis, that the data has Hardy-Weinberg proportions.

Based on this, only the Caucasian-American subjects fit the Hardy-Weinberg equilibrium, whereas the African-American subjects with a Chi-square value of 9.11 do not fit. Some reasons and forces, which lead to an evolutionary change, in which

the Hardy-Weinberg law might fail to apply are: mutation, genetic migration, genetic drift, nonrandom mating and natural selection.

3.3.2 Coriell Cell Repository samples analysis

For the analysis of the human AS3MT 5'-FR, sequences for 120 Coriell Cell Repository samples, 60 for African-American and 60 for Caucasian-American subjects, were analyzed.

The results showed:

- a polymorphism (–114)bp upstream from startcodon ATG, 23 nucleotides upstream of the beginning of the VNTR;
- the WT nucleotide of the SNP had a G, the variant nucleotide showed a C at that position;
- this SNP was linked with a 2 repeat of the VNTR, but was not observed with the allele associated with the 3 or 4 repeat VNTR.

The question was how these differences in the FR would affect the promotor activity and the transcription regulation. Our assumption was that the core promotor of this gene could be located upstream of the VNTR in the FR.

Therefore we analyzed the 120 Coriell samples concerning the SNP in the FR, the number of repeats and the association with the cSNP at nucleotide 860 in exon 9.

A further goal was to design four reporter gene constructs: a 2 repeat VNTR with and without the SNP, one construct including a 3 and one a 4 repeat VNTR.

3.3.2.1 Analysis of 120 Coriell Cell Repository samples concerning the SNP in the FR, the number of repeats and the cSNP variant 860 in exon 9

The arsenic methyltransferase variant 860 displayed 350% of the methylation activity compared with the WT protein. Our next focus became a possible linkage between the cSNP at nucleotide 860, the number of VNTR copies, and the SNP (-114)bp associated with the two repeat VNTR. The SNP upstream results in a change from G to C in the DNA sequence. 60 Coriell Cell Repository DNA samples from Caucasian-Americans and 60 samples from African-Americans subjects were analyzed.

Table 3: Analysis of 120 samples, 60 Caucasian-Americans (CA) and 60 African Americans (AA), for the SNP located 23 nucleotides upstream from the beginning of the VNTR G→ C, VNTR genotype and the SNP at nucleotide 860. X indicates all samples, which have the cSNP in exon 9, T→ C.

CA	SNP (-114)	VNTR repeat	V 860		AA	SNP (-114)	VNTR repeat	V 860
CAU1	G	3,3			AA1	G	2,3	
CAU2	G	3,3			AA2	G	3,3	
CAU3	G	3,3			AA3	G/C	2,3	
CAU4	G	3,3			AA4	G/C	2,3	X
CAU5	G	2,3			AA5	G	2,3	
CAU6	G	2,3			AA6	G	3,4	
CAU7	G/C	2,3	X		AA7	G	2,3	
CAU8	G/C	2,2	X		AA8	G/C	2,3	
CAU9	G	3,3			AA9	G/C	2,3	
CAU10	G	3,3			AA10	G/C	2,3	X
CAU11	G	3,3			AA11	G/C	2,4	
CAU12	G	3,3			AA12	G	3,3	
CAU13	G	2,3			AA13	G/C	2,3	X
CAU14	G	3,3			AA14	G/C	2,3	
CAU15	G	2,3			AA15	G	3,3	
CAU16	G	3,3			AA16	G	3,3	
CAU17	G	2,3			AA17	G	3,3	
CAU18	G/C	2,3	X		AA18	C	2,2	X
CAU19	G	2,2	X		AA19	G/C	2,4	X
CAU20	G/C	2,3	X		AA20	G	3,3	
CAU21	G	2,2			AA21	G	3,3	
CAU22	G	3,3			AA22	G	3,3	
CAU23	G	3,3			AA23	G/C	2,4	X
CAU24	G	2,3			AA24	G/C	2,3	X
CAU25	G	2,3			AA25	C	2,2	X
CAU26	G	3,3			AA26	G	2,3	
CAU27	G	3,3			AA27	G/C	2,4	X
CAU28	G/C	2,2	X		AA28	G/C	2,3	
CAU29	G	3,3			AA29	G	3,3	
CAU30	G/C	2,2	X		AA30	G	3,3	
CAU31	G	3,3			AA31	G	3,3	
CAU32	G	2,3			AA32	G	2,3	
CAU33	G	3,3			AA33	G	3,3	
CAU34	G/C	2,3	X		AA34	G	2,4	
CAU35	G	3,3			AA35	G/C	2,3	
CAU36	G	2,2			AA36	G	2,3	
CAU37	G	3,3			AA37	G	2,3	
CAU38	G	3,3			AA38	G/C	2,2	
CAU39	G/C	2,3	X		AA39	G	2,3	
CAU40	G/C	2,3	X		AA40	G/C	2,4	X
CAU41	G/C	2,3			AA41	G	2,3	
CAU42	G	2,2			AA42	G	3,3	
CAU43	G/C	2,2	X		AA43	G	2,2	
CAU44	G	2,3			AA44	G	3,3	
CAU45	G	2,2			AA45	G/C	2,3	X
CAU46	G	3,3			AA46	G	3,3	
CAU47	G	2,2			AA47	G/C	2,3	
CAU48	G	2,3			AA48	G/C	2,2	X
CAU49	G	2,2			AA49	G	3,3	
CAU50	G	3,3			AA50	C	2,2	
CAU51	G	2,3			AA51	G	2,3	
CAU52	G	2,3			AA52	G/C	2,4	
CAU53	G	2,3			AA53	G	3,3	
CAU54	G	2,3			AA54	G	2,3	X
CAU55	G	2,3			AA55	G/C	2,3	X
CAU56	G/C	2,3	X		AA56	G/C	2,3	X
CAU57	G	3,3			AA57	G/C	2,3	X
CAU58	G	3,3			AA58	G	3,3	
CAU59	G	3,3			AA59	G	3,3	
CAU60	G	2,3			AA60	G	3,3	

The analysis showed a significant linkage between the cSNP at nucleotide position 860 and not only the 2 repeat VNTR, but also the SNP in the FR that was linked to the 2 repeat VNTR. Out of 120 analyzed samples, 28 samples were associated with the cSNP in exon 9:

- 7 samples were homozygous for the two repeat, at least one of the repeats showed the SNP upstream in the FR;
- 15 samples were heterozygous for a two and three VNTR repeat, the two repeat with the SNP in the FR;
- 4 samples were a two and four repeat, the two repeat also associated with the SNP;
- 2 out of the 28 samples with the cSNP in exon 9, at nucleotide position 860, had no linkage between the SNP associated with the two repeat VNTR and the cSNP 860.

3.4 Designing constructs

3.4.1 DNA sequence

DNA sequences within the 5'-UTR and 5'-FR of the human AS3MT gene were amplified to analyze the promotor activity of this gene. The amplified sequences enclosed a 2, 3, or 4 repeat copy of the VNTR, located within this region. The VNTR consists of 36 and 35 nucleotide repetitive sequences, the last copy of the repetitive elements has 35 nucleotides. The VNTR starts in the FR and ends in intron 1.

Table 4: Analysis of the AS3MT 5'-FR and UTR.. This table shows the sequence which was amplified for the analysis of the AS3MT 5'-FR and UTR. The VNTR, here a 3 repeat, bold letters, starts in the FR and ends in intron 1. The underlined and bold sequence is exon 1, 78 bp. The capital A is the first nucleotide of the start codon. The SNP is located (-114), G/C.
The 36 nucleotide repetitive sequence for the VNTR is: GTC CGA GGC CGA GGA GAC AGT GAG TGC GCG CCC TGA, the last copy of the repeat in a gene consists of 35 bp, the last nucleotide "A" is missing. The forward primer hybridizes (-312)bp in the 5'-FR and the reverse primer hybridizes in intron 1.

```
                 Forward Primer
            ACGAGATTTATCCGTGAAAATCGCAGTGGAGCGAACAAGGGGGATGCTGATACCGACCTC
          1 ---------+---------+---------+---------+---------+---------+
   60       TGCTCTAAATAGGCACTTTTAGCGTCACCTCGCTTGTTCCCCCTACGACTATGGCTGGAG


            CTGGTTGGAAAGCCTGTAGAGCAGCGCGGATGACAGTGGAACAGCGGGTCAGGCGGTGGG
         61 ---------+---------+---------+---------+---------+---------+
  120       GACCAACCTTTCGGACATCTCGTCGCGCCTACTGTCACCTTGTCGCCCAGTCCGCCACCC


            CGCCTGAGCGAGGGGCTAGAGCGGGATGGGCGGGCGGAGCAAGCCTGCCAGCCTGGGCGG
        121 ---------+---------+---------+---------+---------+---------+
  180       GCGGACTCGCTCCCCGATCTCGCCCTACCCGCCCGCCTCGTTCGGACGGTCGGACCCGCC


                                             C
            GGCCTCGGCACAGGAGCTGGCTGCGGGAGCCCGCCGTCCTGAGTCGCAGGCCGAGGAGAC
        181 ---------+---------+---------+---------+---------+---------+
  240       CCGGAGCCGTGTCCTCGACCGACGCCCTCGGGCGGCAGGACTCAGCGTCCGGCTCCTCTG
                                             G


            AGTGAGTGCGCGCCCTGAGTCGCAGGCCGAGGAGACAGTGAGTGCGCGCCCTGAGTCGCA
        241 ---------+---------+---------+---------+---------+---------+
  300       TCACTCACGCGCGGGACTCAGCGTCCGGCTCCTCTGTCACTCACGCGCGGGACTCAGCGT


            GGCCGAGGAGACAGTGAGTGCGCGCCCTGGGCGCCCCGCCCCAGCCCCCAGCCCCTTCCCT
        301 ---------+---------+---------+---------+---------+---------+
            CCGGCTCCTCTGTCACTCACGCGCGGGACCCGCGGGGCGGGGTCGGGGGTCGGGGAAGGGA
  360                                                    Reverse Primer
```

3.4.2 Designing pCR2.1 constructs

For the analysis of the promotor activity, a sequence which enclosed the 5'-UTR VNTR, as well as part of the 5'-FR, was amplified with PCR. Two primers were designed, one in forward direction, (-312)bp in the 5'-FR and a reverse primer in intron 1, 48 bp. Both primers contained restriction sites. The forward primer had a restriction site for the enzyme Acc65I, and the reverse primer for the restriction enzyme XhoI. A series of Coriell samples was chosen to amplify repeats and the SNP associated with the 2 repeat VNTR. To test the repeats of different lengths as well as amplification, amplicons were subjected to electrophoresis. Sizes of the amplicons varied between 325 and 397 bp, depending on the number of the repeats in the VNTR. After successful amplification, it was necessary to add an A overhang, for subcloning, because the vector pCR2.1 has a T overhang.
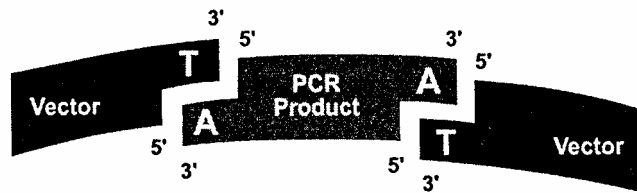

Fig. 8:Vector pCR2.1 overhang: the vector pCR2.1 has a T overhang, so an A must be added to the PCR product for subcloning.

The ligation of the insert into the vector was performed as described in the Methods section (2.11.2). Six samples were transformed for each of the 4 constructs.

The constructs included:

- a 2 repeat VNTR with the SNP at (–114)bp G to C;
- a 2 repeat VNTR with the WT nucleotide G at position (–114)bp;
- the 3 and 4 repeat VNTR with the WT nucleotide G at position (-114)bp.

These constructs were transformed into E.coli INVαF cells. The bacterial colonies were grown on agar plates with β-Gal and ampicillin. After 24 h, blue and white colonies could be distinguished on the plates. Since pCR2.1 is ampicillin resistant, only bacterial colonies containing the insert were able to replicate on the agar.

The pCR2.1 vector contains a galactosidase gene, which makes it possible to differentiate between bacterial colonies that contain the subcloned insert in the vector and those with an empty vector. Colonies containing an insert have a white color because the insert interrupts the galactosidase gene and X-Gal cannot function as a substrate. The white colonies were picked and grown overnight in LB-media. Plasmids containing inserts were purified from the bacterial cells with the Qiagen Plasmid Mini Kit (2.7.2).

A total of 30 samples were sent to the Mayo Molecular Biology core facilities for sequencing and the primers M13 Forward and Reverse (Qiagen) were used to sequence the insert in the vector. No additional polymorphisms were found during analysis of the sequences and no sequence changes had been introduced upon amplification.
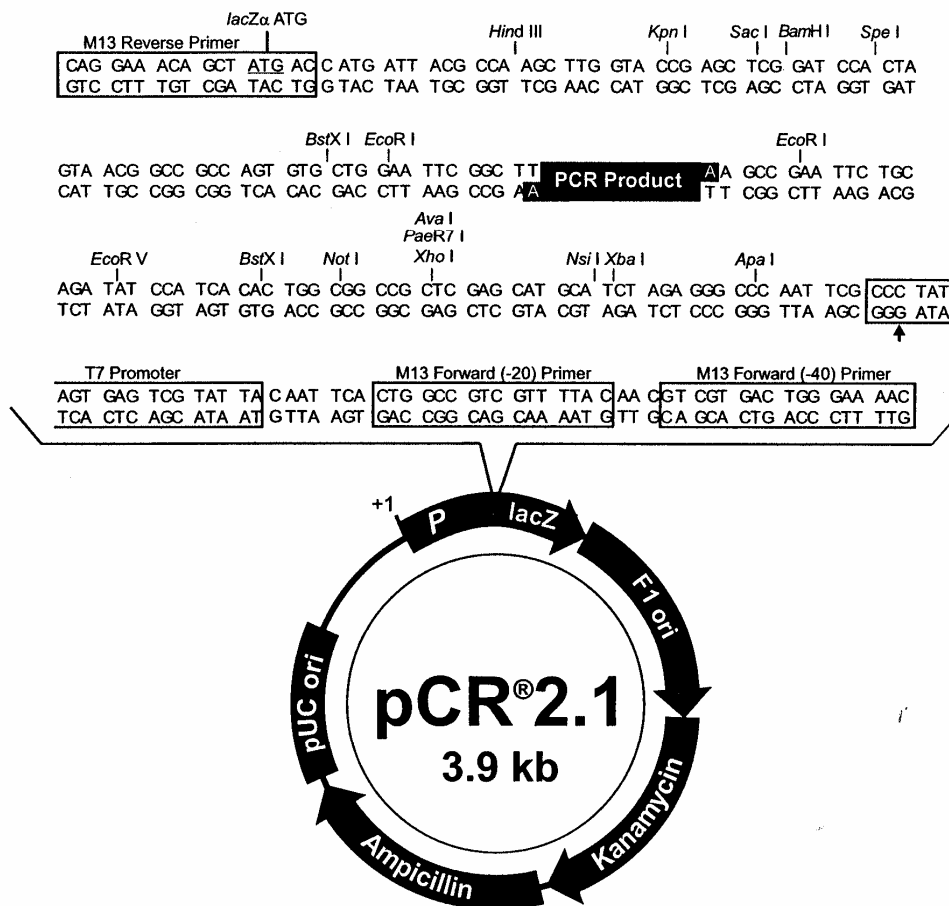
Fig. 9: Vector pCR2.1 map: its sequence, restriction sites and vector primers.

3.4.3 Designing pGL-3 constructs

The pCR2.1 and pGL-3 vectors were digested with restriction enzymes XhoI and Acc65I. Since the primers used to amplify the 5'-FR and UTR were designed with those restriction sites, the inserts in pCR2.1 also contained those restriction sites. The vector pGL-3 itself also has Acc65I and XhoI restriction sites. The vectors were cut at those specific restriction sequences.

Four out of the 30 samples subcloned into pCR2.1 were selected, based on the quality of the sequence, and digested:

- a 2 VNTR repeat;
- a 3 VNTR repeat;
- a 4 VNTR repeat;
- a 2 VNTR repeat linked to the variant SNP(-114)bp.

Digested products were then cloned into pGL-3 basic upstream of the firefly luciferase ORF to create the reporter gene constructs. Each of the inserts had been sequenced on both strands to assure DNA sequence fidelity.

The bands were visualized after electrophoresis in a gel that contained less than 0.8% agarose. Gel bands were excised with a Geneclean Kit. One of the pGL-3 bands was combined with a band of one of the four different inserts that had been digested from pCR2.1, and was ligated overnight at 14°C.

The pGL-3 constructs were transformed into E.coli DH5α cells and plated on agar with ampicillin added. pGL-3 includes no gene which would facilitate distinguishing between colonies with and without inserts. These colonies were picked and grown in LB media with ampicillin. The colonies were screened with PCR to amplify possible inserts subcloned in the vector using the same primers as previously, F(-312)FR and R I1(48).

3.5 Luciferase assay

The luciferase reporter gene constructs were generated to study the regulation of the human AS3MT transcription. Since AS3MT is expressed in kidney and liver, we used

HepG2 and HEK293 cells to express these constructs. 2 μg of purified plasmid DNA was transfected into HepG2 and HEK293 cells together with 200 ng of pRL-TK DNA. The Renilla luciferase activity expressed by pRL-TK was used as a control for transfection efficiency. Cells were also transfected (2.11.3) with pGL-3 basic that lacked the insert, incubated at 37°C for 48 h and harvested.

Each transfection was performed in triplicates, and all transfections were repeated four times to assure reproducibility. The Dual-Luciferase Reporter assay (2.15.2) was used to measure luciferase activity in the transfected cell lysates. Specifically, 20 μl cell lysate was mixed with 100 μl LAR that contained Firefly luciferase. Renilla luciferase activity was measured in the same tube after the addition of 100 μl Stop&Glo reagent. Results were reported as the ratio of Firefly luciferase light units to Renilla luciferase.

The ability of these human AS3MT constructs to drive luciferase expression is listed in the subsequent tables. In both cell lines, luciferase activity was normalized to that observed after transfection with the basic construct. The results demonstrated a cell line-dependent variation in transcription and differences in luciferase activity among constructs with different VNTR lengths. The overall level of activity was similar in HEK293 and HepG2 cells. Constructs that included the AS3MT 5'-FR and 5'-UTR sequence showed significant increases in luciferase activity in both cell lines over that observed in the presence of the "basic" construct.

3.5.1 Results of the luciferase assay

In <u>HEK</u> cells, enhanced transcription was associated with the increasing number of repeats (Fig. 10). Furthermore, the SNP at nucleotide (-114)bp in linkage with the allele of the 2 repeat did not significantly influence the transcription.
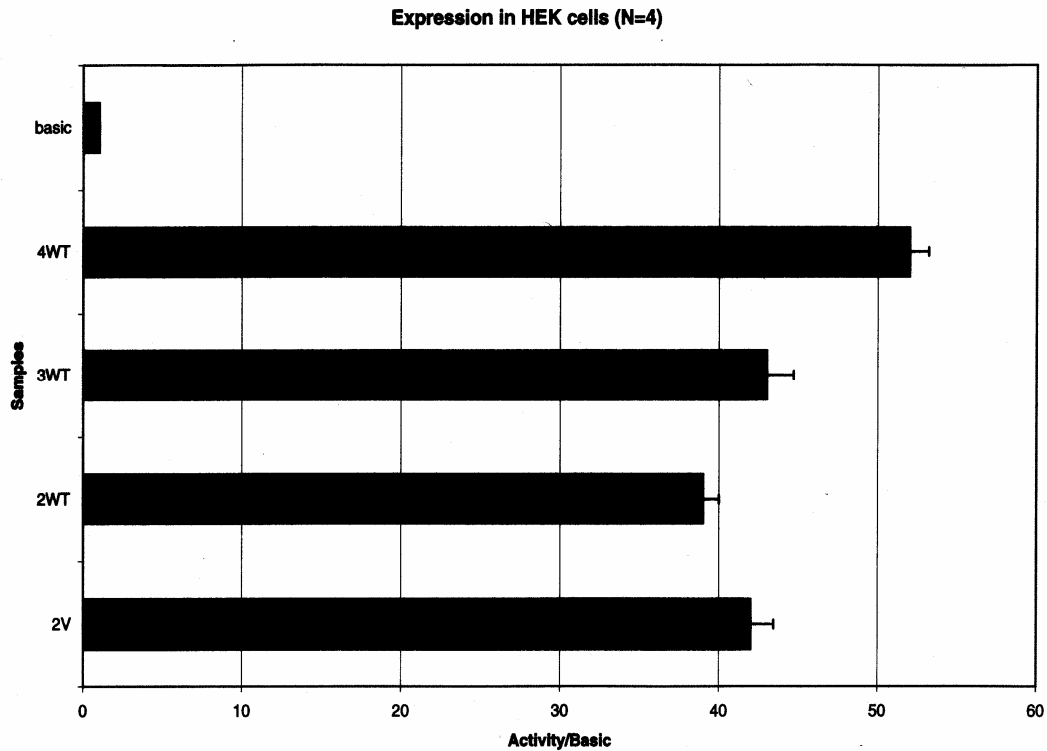

Fig. 10: Average levels of Firefly luciferase activity for AS3MT VNTR reporter constructs transfected into HEK cells. The constructs are labeled with a number, which indicate the copies of the repeat. The number are labeled with WT for wildtype and V for variant . The variant has a "C" instead of the nucleotide "G" at the position (-114)bp. The "basic" displays the activity of an empty pGL-3 vector, values have been corrected for transfection efficiency (activity/basic). Bars represent the average of 4 independent transfections (N=4).

In <u>HepG2</u> cells, however, enhanced transcription was associated with the decreasing number of repeats (Fig. 11). Furthermore, the SNP at nucleotide (-114)bp in linkage with the allele of the 2 repeat (2V) showed little influence on the transcription.
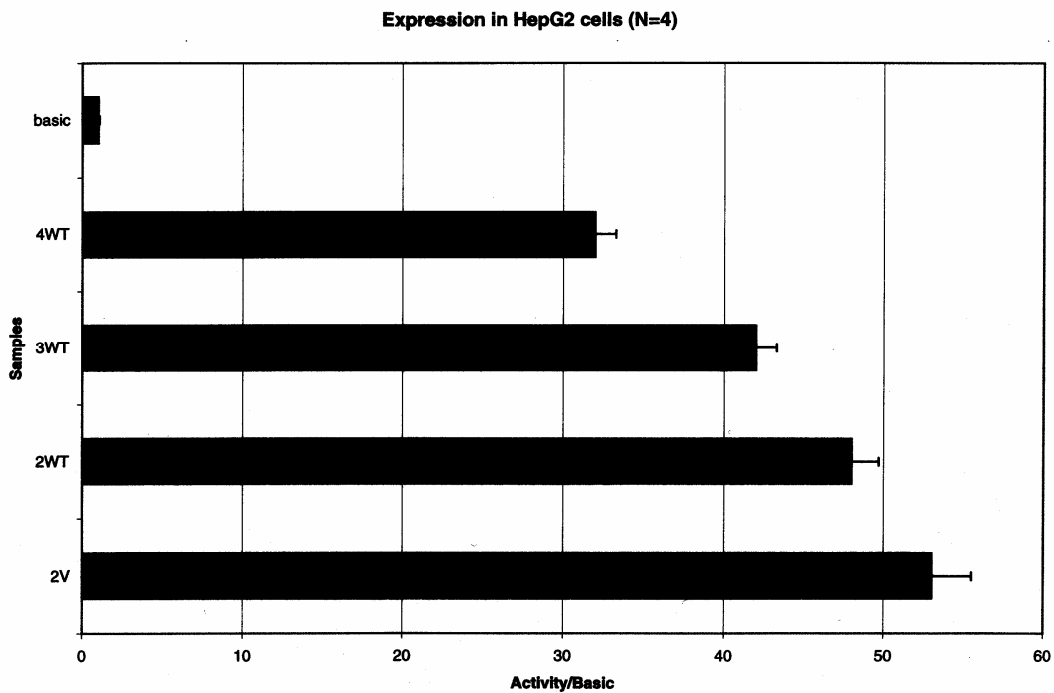
Expression in HepG2 cells (N=4)

Fig. 11: Average levels of Firefly luciferase activity for AS3MT VNTR reporter constructs transfected into HepG2 cells. The constructs are labeled with number, which indicate the copies of the repeat. The numbers are labeled with WT for wildtype and V for variant . The variant has a "C" instead of the nucleotide "G" at the position (-114)bp. The "basic" displays the activity of an empty pGL-3 vector, values have been corrected for transfection efficiency (activity/basic). Bars represent the average of 4 independent transfections (N=4).

Since all constructs were able to drive luciferase expression in the both cell lines, the next interest focus was the core of the promotor.

The aim was to distinguish which portion of the construct might have the most influence on reporter gene activity. Our goal was to design five additional constructs; two in the FR, one containing the SNP at (-114)bp and one with the WT sequence (the nucleotide G), at that position. The other three constructs would contain the 2, 3 and 4 VNTR repeat.

To amplify sequences for the two constructs in the FR, the same forward primer, (-312)bp was used, and a new reverse primer was designed to start in the FR at position (-91)bp. For the VNTR repeat constructs, the forward primer was located (-90)bp in the FR and the reverse primer intron I (16)bp. When we performed electrophoresis with the amplicons and then PCR, it became clear, that the VNTR repeats did not amplify, although several different conditions were tried. The conclusion was that the presence of a CpG island and/or the repeat itself made it very

difficult to amplify this region. A CpG island is a DNA sequence which consists of many C and G nucleotides. For a successful amplification a new set of primers, hybridizing both upstream and downstream from the VNTR was used. If we changed only one of the primers sitting directly at the ends of the VNTR, there was still no amplification. Our idea was, that the amplicons were too small to be seen on the electrophoresis gel. Their seizes differed between 122-220 bp.

Therefore, we started to subclone the assumed constructs into vector pCR2.1 and sequenced on both strands. After the digestion of pCR2.1 containing the inserts, and the cutting of pGL-3 with the restriction enzymes Acc65I and XhoI, five inserts were ligated into pGl-3: the 2, 3, 4 repeat and the FR with and without the SNP (-114)bp. The purification was repeated several times with one of two methods, gel purification or PCR. The transformation was also repeated more than once. We transformed into E.coli TopTen, and DH5α cells. All attempts failed to give subcloned inserts in pGl-3. We then used another reverse primer while screening the colonies to increase the seize of the amplicons:

- the primer M13 sat in the vector directly behind the insert;
- we also designed a new primer, pGL-3 rev402, to increase the seize of the amplicons significantly;
- the Maxiprep for the vector pGL-3 was also repeated.

However, all screenings for these inserts in the vector pGL-3 failed.

4. Discussion

"Pharmacogenetics to come"…."Genetically selected medicine has been much hyped but has significant potential. Regulation and treatment will depend on pharmaceutical companies more readily sharing genetic data" (Goetz et al., 2004).

Seeing the complications of drug therapy like hundred of deaths worldwide on the one hand and or no response at all while giving the same drug dose to patients on the other hand, this is a tremendous step into the future medicine. SNPs in genes are responsible for such diversity of the side effects. The future might be the personalized medicine according to the analyzed genes and physicians can expect to receive more pharmacogenetic data to help choosing the right drug therapy and dose for the individual patient.

4.1 AS3MT and the pioneer of pharmacogenetics

The AS3MT gene encodes for an enzyme that catalyzes the methylation of arsenic. After resequencing the human AS3MT gene, functionally significant nonsynonymous cSNPs in the ORF, moreover in the 5-FR and UTR were identified. We tested the hypothesis how these polymorphisms might contribute to individual differences in arsenic methylation and in which way the polymorphisms may effect in vivo.

Seeing the enzyme TPMT as a pioneer model of the pharmacogenetic progress, we followed the idea that we would find the answer to the tremendous differences in the methylating activity of arsenic in the SNPs of the AS3MT gene. We proposed that the cSNPs, found in the ORF, 5'-FR and UTR, were responsible for the individual differences in the methylation level of arsenic.

4.2 Expression of the cSNPs in the ORF

Functional genomic studies of the AS3MT WT sequence and the 3 nonsynonymous SNPs in the ORF were performed after transfection in a mammalian cell line, COS-1. Although we first struggled during preliminary studies of the enzyme assay and our

counts were variable, we found a methylation activity pattern, which was later proved during optimized studies.

We found that two of the three cSNPs within the ORF of the AS3MT, resulted in a protein with reduced enzymatic activity compared to the WT:

- variant 517 (exon 6);
- variant 917 (exon 10);
- variant 917 (exon 10) linked to variant 860 (exon 9).

The activity of the recombinant variant 860 allozyme seemed to be higher than of the WT allozyme. The variant 860 has in its sequence a cSNP in exon 9, with T instead of C. Many CpG (cytidine-phosphatidyl-guanosine) dinucleotides are methylated on the cytosine base, and spontaneous deamination of methyl-C residues can give rise to T residues. As a result, methyl-CpG dinucleotides steadily mutate to TpG (thymidine-phsphatidyl-guanosine) dinucleotides (International Human Genome Sequencing Consortium, 2001).

The higher methylation activity of the variant 860, which has a frequency of 10% in African-American and 10% in Caucasian-American subjects, raises the possibility that the nucleotide sequence for the variant 860 could be the nucleotide sequence for the ancestral WT, the C being methylated during evolution.

4.3 Influence of the polymorphisms in the VNTR region on the AS3MT gene activity

To define the possible contribution of gene sequence variation to the regulation of the AS3MT transcription, we designed four reporter gene constructs. The constructs consisted of the 4, 3 and the 2 repeats for the 5'-UTR VNTR, the 2 repeats with the WT, G, and the other with the SNP, C, at position (-114)bp in the FR.

After subcloning of the inserts into pGL-3 upstream of the Firefly luciferase ORF, we assayed the luciferase activity with a Dual-Luciferase Reporter assay system.

The reporter gene constructs were transfected into HepG2 and HEK293 cells to compare the ability of the constructs to drive luciferase expression. HepG2 and

HEK293 cells displayed the same level of activity when transfected with the constructs.

Surprisingly, we found that:

- differences in activity between the constructs in the two cell lines seemed to be opposite in direction, especially, the two repeat associated with and without the SNP gave the highest activity in the HepG2 liver cell line, whereas, they had lower activity in the HEK293 kidney cell line.

The same occurred with the four repeat VNTR:

- in HepG2 cells, the more repeats present, the less luciferase activity we found, however HEK293 cells showed the highest activity for the four repeat VNTR.

In summary, we observed differences in luciferase activity for the same constructs expressed in the two cell lines. These observations suggested that AS3MT transcription differs between HepG2 and HEK293 cells transfected with the same constructs. This indicates cell-line dependent transcription. In the future it will be interesting to analyze the transcription factors involved.

4.4 Linkage between the 2 VNTR repeat and the variant 860

All 28 samples with the cSNP at the nucleotide position 860 were linked with the 2 repeat VNTR. In addition 26 out of the 28 cSNP 860 samples had at least one allele that was linked to the 2 repeat VNTR associated with the SNP(-114)bp upstream. It seemed that the higher gene activity in the HepG2 cells due to the 2 VNTR repeat associated with the SNP in the promotor region produced a higher activity level.

4.5 Conclusions and future approaches

In conclusion, we were able to prove that the differences in the methylation level are due to the polymorphisms in the AS3MT gene. Furthermore, we showed that not only the ORF but also the region within the FR and UTR had an impact on the enzyme activity.

Two cSNPs within the ORF of the AS3MT gene, resulted in a protein with reduced enzymatic activity, variant 517 (exon 6) and variant 917 (exon 10). Western analysis showed that these two allozymes had not more than 20% of the level of WT immunoreactive protein (variant 517: 20%, variant 917: 4% and the "double allozyme" 860/917 8%).

These results fit several reports, which show that the alteration of one or two amino acids as a result of common polymorphisms within a gene, can result in a protein degradation through proteasome-mediated process. Moreover, recent results indicate that chaperone proteins seem to play an important role in the process of protein degradation (Wang et al., 2003).

A common polymorphism (*3A) of the TPMT gene has two alterations in its amino acid sequence and is much more rapidly degraded than the WT allozyme of TPMT. It appeared that ubiquitination followd by a proteasome-mediated degradation of the variant TPMT allozyme was involved (Weinshilboum et Wang, 2004). Prof. R.M. Weinshilboum and Dr. Wang also presented results, which indicate that the polymorphism *3A of TPMT is much more highly associated with heat shock proteins like hsp 90 and hsp 70 than with the TPMT wildtype. This association with heat shock proteins has influences on ubiquitination and degradation.

Therefore, Prof. R.M. and his coworkers tested protein degradation of AS3MT. This was studied in a rabbit reticulocyte lysate with the *3A TPMT polymorphism as a positive control for rapid degradation. During an observation time of 24 h there was no significant difference observed for the AS3MT allozymes.

Since rapid degradation does not appear to be the mechanism for differences in levels of the observed AS3MT proteins, the specific mechanisms for degradation of the AS3MT allozymes variant 517 and variant 917 remain to be determined in the future.

4.6 Results after my stay

Prof. R.M. Weinshilboum and his coworkers took further steps to analyze the cSNP at the nucleotide position 860 in exon 9. This cSNP had a frequency of 10% in both populations studied, altering the amino acid at position 287 from Met to Thr. In 10 primates exon 9 was also amplified. These primates were: rhesus monkey, pigtailed macaque, bonobo, gorilla, chimpanzee, sumatian orangutan, red-chested mustached tamarin, black-handed spider monkey, common wooly monkey and ring-tailed lemur, all obtained from the Coriell Cell Repository. It appeared that this codon encoded the human variant amino acid Thr in all of the ten primate samples. This implies that the amino acid sequence Met coding for the human WT at the position 287 had occurred very recently in primate evolution. The hypothesis that cSNP at nucleotide position 860 might express the amino acid sequence for the ancestral WT protein could be an explanation. The possible function of natural selection remains to be discussed.

The haplotype analysis (Schaid et al., 2002) had shown that the cSNP at nucleotide position 917 (Ile306) was only observed in the presence of the cSNP at nucleotide position 860 (Thr287). During my preliminary studies I often failed in the methylation activity assay to detect methylation activity for the allozyme variant 917. Since this variant exists only as a "double variant" in nature, it is not surprising that the methylation activity in the assays was difficult to detect.

4.7 Clinical approach

To take the step from the laboratory research to the bedside medicine:
The pharmacogenetic model TPMT shows the future direction of the AS3MT research. TPMT has an influence on the treatment with 6-MP against leukemia. The cytotoxic effects of 6-MP are due to thioguaninnucleotides. 6-MP is a so called "prodrug". When patients are undertreated with 6-MP, less thioguanine was produced, or overtreated, more thioguanine resulted from the treatment with 6-MP. With undertreatment leukemia cells survived, with overtreatment the cytotoxic effects were too high for the body. It was shown that the activity of the TPMT and amount of thioguanine were linked together in an opposite way. This is due to the methylation of thioguanine in the presence of TPMT (Weinshilboum, 1992).

The striking observation of our studies showed increased levels of both enzyme activity and protein levels for the common variant 860 (Thr287) allozyme when compared to the AS3MT WT. The frequency of this polymorphism is about 10% in both American-African and Caucasian-American subjects, approximately 1% of the subjects in both populations would be homozygous for this allele and might express increased levels of AS3MT and perhaps elevated arsenic methylation. Those subjects might also be at a higher risk for cytoxic, genotoxic effects of arsenic exposure and the therapy with arsenic containing drugs, if our assumption is the case, that methylation increases arsenic toxicity.

These examples show possible impacts and mechanisms of the pharmacogenetics of TPMT and AS3MT genes on medicine. Given the potential value of knowing all the possible factors that influence the effects of new agents, it is likely that pharmacogenetics will have an increasingly important role in drug discovery and development (Evans and Relling, 2004).

The AS3MT will go into the same direction as its model TPMT, the initial steps are done by resequencing and analyzing the polymorphisms of the gene. A major challenge for the future will be finding ways to translate the pharmacogenetic information into meaningful clinical reality and individual patient treatment.

5. References

Aletta JM, Cimato TR, and Ettinger MJ (1998): Protein methylation: a signal event in post-translational modification. *Trends Biochem Sci 23: 89-91*

Bertsche B, and Schulz M (2003): Arsentrioxid zur Leukaemiebehandlung. *Pharm Ztg 148. Jahrgang: 3289-3294*

Brookes AJ (1999): The essence of SNPs. *Gene 234: 177-186*

Buchet JP, and Lauwerys R (1988): Role of thiols in the in-vitro methylation of inorganic arsenic by rat liver cytosol. *Biochem Pharmacol 37: 3149-3153*

Cairns J (1975): Mutation selection and the natural history of cancer. *Nature 255: 197-200*

Cantoni GL, and Durell J (1957): Activation of methionine for transmethylation II. The methionine-activation enzyme; studies on the mechanism of the reaction. *J Biol Chem 225: 1033-1048*

Chouchane S, and Snow ET (2001): In vitro effect of arsenicals compounds on glutathione-related enzymes. *Chem Res Toxicol 14: 517-522*

Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, and Walters L (1998): New goals for the U.S. Human Genome Project: 1998-2003. *Science 282: 682-689*

Del Razo LM, Styblo M, Cullen WR, and Thomas DJ (2001): Determination of trivalent methylated arsenicals in biological matrices. *Toxicol Appl Pharmacol 174: 282-293*

Enzyme Nomenclature (1992): Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes. *New York: Academic Press*

Evans WE, and Relling MV (2004): Moving towards individualized medicine with pharmacogenomics. *Nature 429: 464-468*

Goetz MP, Ames MM, and Weinshilboum RM (2004): Primer on medical genomics. PartXII: Pharmacogenomics--general principles with cancer as a model. *Mayo Clin Proc 79: 376-384*

Gordon D, Abajian C, and Green P (1998): Consed: a graphical tool for sequence finishing. *Genome Res 8: 195-202*

Healy SM, Casarez EA, Ayala-Fierro F, and Aposhian H (1998): Enzymatic methylation of arsenic compounds. V. Arsenite methyltransferase activity in tissues of mice. *Toxicol Appl Pharmacol 148: 65-70*

Holliday R, and Grigg GW (1993): DNA methylation and mutation. *Mutat Res 285: 61-67*

International Agency for Research on Cancer (1980): IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Metal and Metallic Compounds. *International Agency for Research on Cancer Vol 23: 39-141*

International Human Genome Sequencing Consortium (2001): Initial sequencing and anlysis of the human genome. *Nature 409: 860-921*

Kaufman RJ (1985): Identification of the components necessary for adenovirus translational control and their utilization in cDNA expression vectors. *Proc Natl Acad Sci USA 82: 689-693*

Kitchin KT, Del Razo LM, Brown JL, Anderson WL, and Kenyon EM (1999): An integrated pharmacokinetic and pharmacodynamic study of arsenite action. 1. Heme oxygenase induction in rats. *Teratog Carcinog Mutagen 19: 385-402*

Klein CB, and Costa M (1997): DNA methylation, heterochromatin and epigenetic carcinogens. *Mutat Res 386: 163-180*

Kruglyak L, and Nickerson DA (2001): Variation is the spice of life. *Nat Genet 27: 234-236*

Lai MS, Hsueh YM, Chen CJ, Shyu MP, Chen SY, Kuo TL, Wu MM, and Tai TY (1994): Ingested inorganic arsenic and prevalence of diabetes mellitus. *Am J Epidemiol 139: 484-492*

Li J, Waters SB, Drobna Z, Devesa V, Styblo M, and Thomas DJ (2005): Arsenic (+3 oxidation state) methyltransferase and the inorganic arsenic methylation phenotype. *Toxicol Appl Pharmacol 204: 164-169*

Lin S, Shi Q, Nix FB, Styblo M, Beck MA, Herbin-Davis KM, Hall LL, Simeonsson JB, and Thomas DJ (2002): A novel S-adenosyl-L-methionine: arsenic (III) methyltransferase from rat liver cytosol. *J Biol Chem 277: 10795-10803*

Masood E (1999): As consortium plans free SNP map of human genome. *Nature 398: 545-546*

Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E, and Sing CF (1998): DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nat Genet 19: 233-240*

Nordstrom DK (2002): Public health. Worldwide occurences of arsenic in ground water. *Science 296: 2143-2145*

National Research Council (NRC) (1999): Arsenic in Drinking Water. *National Academy Press, Washington D.C. 1999*

National Research Council (NRC) (2001): Arsenic in Drinking Water. *National Academy Press, Washington, D.C. 2001 Update*

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cock L, Davenprtz R, Dante M, Fulton L, Hillier L, Waterston RH, Mc Pherson JD, Gilman B, Schaffners S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D, and International SNP Map Working Group (2001): A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature 409: 928-933*

Schaid DJ, Rowland CM, Tines DE, Jacobson RM, and Poland GA (2002): Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet 70: 425-434*

Smith AH, Lopipero PA, Bates MN, and Steinmaus CM (2002): Public health. Arsenic epidemiology and drinking water standards. *Science 296: 2145-2146*

Styblo M, Delnomdedieu M, and Thomas DJ (1996): Mono- and dimethylation of arsenic in rat liver cytosol in vitro. *Chem Biol Interact 99: 147-164*

Styblo M, Serves SV, Cullen WR, and Thomas DJ (1997): Comparative inhibition of yeast glutathione reductase by arsenicals and arsenothiols. *Chem Res Toxicol 10: 27-33*

Thigpen KG (2004). International sequencing consortium: *Environ Health Perspect 112: A406*

Thompson DJ (1993): A chemical hypothesis for arsenic methylation in mammals. *Chem Biol Interact 88: 89-114*

Tollervey D (1996): Small nucleolar RNAs guide ribosomal RNA methylation. *Science 273*: *1056-1057*

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

Weinshilboum RM (1988): Pharmacogenetics of methylation: relationship to drug metabolism. *Clin Biochem 21: 201-210*

Weinshilboum RM (1992): Methylation pharmacogenetics: thiopurine methyltransferases as a model system. *Xenobiotica 22: 1055-1071*

Weinshilboum RM (2001): Thiopurine pharmacogenetics: clinical and molecular studies of thiopurine methyltransferases. *Drug Metab Dispos 29: 601-605*

Weinshilboum RM (2003): Inheritance and drug response. *N Engl J Med 2003 348: 529-537*

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

Wood TC, Salavagionne OE, Mukherjee B, Wang L, <u>Klumpp AF</u>, Thomae B, Eckloff BW, Schaid DJ, Wieben ED, Weinshilboum RM (2006): Human arsenic methyltransferase (AS3MT), pharmacogenetics: gene resequencing and functional genomics studies. *J Biol Chem 281: 7364-7373*

World Health Organization (1981): Environmental health criteria 18: Arsenic Geneva

Zakharyan R, Wu Y, Bogdan GM, and Aposhian HV (1995): Enzymatic methylation of arsenic compounds: assay, partial purification, and properties of arsenite methyltransferases and monomethylmethylarsonic acid methyltransferase of rabbit liver. *Chem Res Toxicol 8: 1029-1038*

Akademische Lehrer

Prof. Dr. Dr. J. Krieglstein

Institut für Pharmakologie und Toxikologie, Philipps-Universität Marburg

Ketzerbach 63

35032 Marburg

Prof. Dr. R.M. Weinshilboum,

Division of Clinical Pharmacology

Department of Molecuar Pharmacology and Experimental Therapeutics,

Mayo Medical School-Mayo Clinic-Mayo Foundation

200 First Street SW

Rochester, MN 55905

USA

Akademische Lehrer

Acknowledgements

At this point of my work, I would like to thank the following individuals without whose help, support, understanding and fond advices this work would not have been accomplished.

"Special places are moveable feasts, and will always stay with you" (Hemmingway). I want to thank Prof. Dr. Richard Martin Weinshilboum from the Department of Molecular Pharmacology and Experimental Therapeutics at the Mayo Clinic Rochester, MN, USA, for the opportunity to work in his laboratory and receiving his continuous support and advices during my stays. I also would like to thank my american coworkers in the laboratory for their help and the introduction to science. Special thanks go to my coworker Ezequiel, who showed me that science is not about following a protocol, but being creative and never giving up.

"Leider läßt sich eine wahrhafte Dankbarkeit nicht mit Worten ausdrücken"(Goethe). Dennoch möchte ich Herrn Prof. Dr. Dr. Josef Krieglstein vom Institut für Pharmakologie und Toxikologie, Philipps-Universität Marburg ganz besonders für eine großartige Unterstützung zur Verwirklichung dieser Arbeit danken.

"There is no place like home" (Die Zauberer von Oz). Meiner Familie; besonders meinen Eltern möchte ich für ihre stetige Begleitung sowie Unterstützung auf meinen Abenteuern und die Hilfe bei der Verwirklichung meiner Träume danken. Meinem großen Bruder gebührt großer Dank, nicht nur für die Hilfe in technischen Fragen, sondern auch für den brüderlichen Beistand.
Susanne danke ich für eine große Chance und eine stetige, rasant schnelle und objektive Unterstützung über alle möglichen Kommunikationswege.

"Der verlorenste aller Tage ist der, an dem man nicht gelacht hat" (Chamfort). Ich danke meinen Freunden, die mir vor, während und nach meiner Zeit in den USA zur Seite standen, mich in der Ferne unterstützt, aufgemuntert und begleitet haben.

Ehrenwörtliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die dem Fachbereich Medizin Marburg zur Promotionsprüfung eingereichte Arbeit mit dem Titel: "Human arsenic methyltransferase pharmacogentics: functional studies of common polymorphisms and its impact on medicine" im Institut für Pharmakologie and Toxikologie unter Leitung von Prof. Dr. Dr. J. Krieglstein mit Unterstützung von Prof. Dr. R.M. Weinshilboum, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Graduate School-Mayo Clinic, Rochester, MN, USA, ohne sonstige Hilfe selbst durchgeführt, und bei der Abfassung der Arbeit keine anderen als die in der Dissertation aufgeführten Hilfsmittel benutzt habe.

Ich habe bisher an keinem in- und ausländischen Medizinischen Fachbereich ein Gesuch um Zulassung zur Promotion eingereicht, noch die vorliegende oder eine andere Arbeit als Dissertation vorgelegt.

Veröffentlichung

Vorliegende Arbeit wurde im folgenden Publikationsorgan veröffentlicht:

Wood TC, Salavaggione OE, Mukherjee B, Wang L, <u>Klumpp AF,</u> Thomae BA, Eckloff BW, Schaid DJ, Wieben ED, Weinshilboum RM (2006):

Human arsenic methyltransferase (AS3MT) pharmacogenetics:
gene resequencing and functional genomics studies.

J Biol Chem 281: 7364-73

Gräfelfing, den 30.05.2008