Pharmacogenetics of Arylamine *N*-acetyltransferase genes in South African Populations

Cedric J. Werely

Dissertation presented for the degree of Doctor of Philosophy

ĵ - ÅÄÉÃÁÌ8"ÉÏÃÈÅ Í ÉÓÔÒÙQ8 in the Faculty of Medicine and Health Sciences at Stellenbosch University

Promoter: Prof. Paul D. van Helden
Dept. of Biomedical Sciences.

December, 2012

Declaration

I, the undersigned, hereby declare that the work contained therein is my own, original work, and that I have not previously in its entirety or in part submitted it at any university for obtaining any degree qualification.

Signature: ----------

C.J. Werely.

Date: ----------

Summary

Tuberculosis (TB) has been declared a global health emergency by the World Health Organisation, and consequently there is an urgency to develop improved methods of diagnosis and treatment. Despite the current TB epidemic, the disease can be treated effectively using isoniazid (INH) in combination with other antibiotics. However, INH is inactivated in the body by certain drug metabolising enzymes, which may reduce the efficacy of TB treatment. The activity of these drug metabolising enzymes, called NAT, are in turn reduced by nucleotide changes (SNPs) in the gene. These genetic variants (alleles) have been correlated with the rapid- (FA), intermediate- (IA), and slow acetylation (SA) enzymatic activity, and one is therefore able to investigate potential phenotypic effects via genotypic analyses.

We investigated these genetic changes in the *NAT1* and *NAT2* genes in individuals from the local Coloured community (SAC) since this group has one of the highest TB incidences in the country. NAT2 is primarily responsible for the inactivation of INH, whilst NAT1 metabolises para-aminosalicyclic acid (PAS) which is used in the treatment of drug resistant TB.

The NAT2 results indicated that the *NAT2* alleles were not equally represented in three local ethnic groups studied, and subsequently the rapid, intermediate and slow acetylation activity reflected these differences. However, the relative frequency of these variants in the SAC and Caucasian groups were relatively low. These differences require further investigation to determine their overall relevance to the NAT2 activity differences between groups.

In the case of the NAT1 analysis we also observed differences in the relative frequency of various *NAT1* alleles between Caucasian and SAC individuals. However, many of these *NAT1* SNPs and alleles have not as yet been characterised, so effects of these variants are currently unknown. Interestingly, the *NAT1*4* and *NAT1*10* alleles were the most prevalent *NAT1* alleles in both Caucasians and SAC. The *NAT1*4* allele exhibits the rapid

NAT1 activity, whilst the activity of the *NAT1*10* allele is currently subject to ongoing debate. In this respect, the analysis of NAT1 continues to be a topic for ongoing research.

These results, observed for the NAT genes, underscore the importance of doing genetic analyses in local ethnic groups, since these differences may vary significantly between the groups.

Opsomming

Tuberkulose (TB) is deur die Wêreldgesondheidsorganisasie (WGO) tot 'n globale gesondheidsnood verklaar en derhalwe is dit noodsaaklik dat nuwe, verbeterde diagnostiese metodes ontwikkel word, wat tot meer effektiewe behandeling kan lei. Ten spyte van die huidige TB-epidemie, kan die siekte doeltreffend behandel word deur middel van isoniasied (INH), in kombinasie te met ander antibiotika. INH kan egter geïnaktiveer word deur sekere ensieme in die liggaam, met die gevolg dat INH nie meer effektief is nie in die behandeling van TB. Die aktiwiteit van hierdie ensiem, die sogenaamde NAT2 (Arielamien N-asetieltransferase 2) ensiem, word op sy beurt beïnvloed deur sekere nukleotied veranderings (SNPs) in die geen. Hierdie genetiese veranderings gekorreleer met ensiemaktiwiteitsveranderings (geklassifiseer as vinnig (FA) Intermediêr (IA) en stadig (SA)), wat mens in staat stel om potensiële fenotipiese effekte te ondersoek deur middel van genotipiese analise.

Ons het hierdie genetiese veranderings ondersoek in die *NAT1* en *NAT2* gene in individue van die Kleurling-gemeenskap (SAC) omdat díe bevolkingsgroep die hoogste voorkoms van TB in die land het. NAT2 is primêr verantwoordelik vir die inaktivering van INH, terwyl NAT1 para-amienosalisilaat (PAS) inaktiveer, wat gebruik word in die behandeling van midel-weerstandige TB.

Die NAT2 resultate dui daarop dat die allele van die *NAT2* geen nie eweredig verteenwoordig wasin die drie etniese groepe nie en derhalwe word die vinnige (FA), intermediêre (IA) en stadige (SA) ensiemaktiwiteite deur hierdie verskille weerspieël. Hoewel die teenwoordigheid van hierdie variante relatief laag was in die SAC en Koukasiër gemeenskappe, is verdere studies nodig om die omvang van hierdie verskille te bepaal ten onsigte van NAT2 aktiwiteit tussen groepe.

In die geval van die NAT1 analise het ons verskille waargeneem in die voorkoms van verskeie *NAT1* allele tussen Koukasiese en SAC individue. Baie van hierdie *NAT1* SNPs is egter nog nie gekarakteriseer nie, en derhalwe is die effek van hierdie NAT1 variante

onbekend. Die *NAT1*4* en *NAT1*10* allele was die prominentste *NAT1* alleel in beide Koukasiërs en SAC. Die *NAT1*4* is betrokke by vinnige NAT1 aktiwiteit, terwyl die effek van die *NAT1*10* alleel nog onderhewig is aan aktiefwe debat. In hierdie verband, is die studie van NAT1 steeds 'n onderwerp vir toekomstige navorsing.

Hierdie resultate, wat vir die NAT gene waargeneem is, beklemtoon die belangrikheid van verdere genetiese analises in plaaslike etniese groepe, aangesien hierdie verskille beduidend kan wees tussen die verskillende groepe.

To my dearest wife Volene, and our children, thanks for being my inspiration, always….

Acknowledgements

I would like to convey my appreciation and sincere gratitude to the many people who cared for me on this journey. You will never know how much I owe you……

Firstly, to Prof. Paul D. van Helden, my promoter and mentor. Thank you for believing in me, even when I don't….

To colleagues and friends, who like me have passed through the doors of the Department of Biological Sciences; the times were tough, but your camaraderie made it all worthwhile.

Many thanks also to the individuals and funding organisations, for allowing me the privilege to pursue this thing we call research.

To my parents, Conrad James Werely, and Rita Dulcie Simons-Werely, - thank you for indulging my dreams.

To GOD be the Glory……

…..***I shall let the little I have learnt go forth into the day in order that someone better than I may guess the truth, and in his work may prove and rebuke my error. At this I shall rejoice that I was yet a means whereby this truth has come to light***

(Albrecht Dürer, 1471-1528)

List of Abbreviations

A	Adenine nucleotide base
Abs	Absorbance
ADR	Adverse Drug Reaction
AIDS	Acquired Immune Deficiency Syndrome
Ala (or A)	Alanine amino acid
AP-1	Activator Protein 1
Arg (or R)	Arginine amino acid
ATP	Adenosine triphosphate
Asn (or N)	Asparagine amino acid
Asp (or D)	Aspartate amino acid
BCG	Bacillus Calmette Guèrin
bp	base pair
C	Cytosine nucleotide base
CI	Confidence Interval
cm	centimetre
CP-PCR	Confronting-primer PCR
Cys (or C)	Cysteine amino acid
D’	Lewontin’s standardised disequilibrium coefficient
ddNTP	dideoxynucleotide tripshosphate
DIH	Drug-induced Hepatotoxicity
DILI	Drug-induced Liver Injury
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanine triphosphate
dTTP	deoxythymine triphosphate
dNTP	deoxynucleotide triphosphate

DOTs	Directly Observed Treatment, short course
DME	Drug Metabolising Enzyme
EBA	Early Bacteriocidal Activity
EDTA	Ethylene Diamine Tetra-Acetate
EMB	Ethambutol
g	gram
G	Guanine nucleotide base
gDNA	genomic DNA
Gln (or Q)	Glutamine amino acid
Glu (or E)	Glutamate amino acid
Gly (or G)	Glycine amino acid
HIV	Human Immuno-deficiency virus
His (or H)	Histidine amino acid
HPLC	High Pressure Liquid Chromatography
HWE	Hardy-Weinberg Equilibrium
IA	Intermediate Acetylator
Ile (or I)	Isoleucine amino acid
INH	Isoniazid
katG	catalase peroxidise enzyme
kb	kilobase pair
kDa	kiloDalton
kg	kilogram
LD	Linkage Disequilibrium
Leu (or L)	Leucine amino acid
Lys (or K)	Lysine amino acid
LTBI	Latent Tuberculosis Infection
M	Molar
MAF	Minor Allele Frequency
MDR	Multi-drug Resistant
Met (or M)	Methionine amino acid
ml	millilitre
mRNA	messenger RNA

Mtb	*Mycobacterium tuberculosis*
n	number of samples
NAT	*N*-acetyltransferase
NCBI	National Centre for Biotechnology Information
NCE	non-coding element
ng	nanogram
nm	nanometre
ns	non significant
OR	Odds Ratio
r^2	Square of the correlation coefficient between two loci
RA	Rapid/Fast acetylator
PCR	Polymerase Chain Reaction
Phe (or F)	Phenylalanine amino acid
PRO	Procainamide
Pro (or P)	Proline amino acid
PZA	Pyrazinamide
RFLP	Restriction Fragment Length Polymorphism
RIF	Rifampicin
RNA	Ribonucleic acid
rmp	revolutions per minute
SA	Slow Acetylator
SAC	South African Coloured
SAP	Shrimp Alkaline phosphatase
SB	Sodium Borate
Ser (or S)	Serine amino acid
SM	Surrogate marker
SMZ	Sulphamethazine
SNP	Single Nucleotide Polymorphism
SP-1	Stimulatory Protein-1
Taq	*Thermus aquaticus*
TB	Tuberculosis
Thr (or T)	Threonine amino acid

Tm	Melting point temperature
Trp (or W)	Tryptophan amino acid
TSS	Transcription start site
Tyr (or Y)	Tyrosine amino acid
µg	microgram
µl	microlitre
µM	micro Molar
UTR	Untranslated region
v	volts
Val (or V)	Valine amino acid
vs	versus
WHO	World Health Organisation
WT	wild type
XDR	Extensively-drug resistant
%	percentage
/	per
°C	Degrees Celcius
5’	5-prime end
3’	3-prime end

Table of Contents

CHAPTER 1

Introduction

1.1 General background to Tuberculosis

Tuberculosis (TB) is the principal cause of death in adults, when seen in the context of the global health of individuals. Reports of the World Health Organisation (WHO) suggest that no less that one-third of the world's population is host to or have been exposed to the causative agent, *Mycobacterium tuberculosis* (*Mtb).* Subsequently it is projected that at least 8 million of these infected individuals will develop active TB, resulting in an annual death rate in excess of 2 million individuals (WHO Reports, 2006 onwards). This is all the more disconcerting since TB is a curable disease, and effective anti-tuberculous treatment has been available for more than 40 years [1]. However, due to several factors, such as social stigma and poor patient compliance, and a lack of investment in a concerted TB control programme, the pandemic has been a global health emergency for more than a decade (WHO Report, 2007). The combined effect of these and other factors has also led to the increasing emergence of multi-drug resistant (MDR), and even more recently, extensively drug resistant (XDR) *Mtb* strains [2]. Therefore, it would appear that the pandemic is still escalating, and in some regions of the world, most notably sub-Saharan Africa and South East Asia, it is fuelled by a synergy with the Human Immunodeficiency Virus and Acquired Immune Deficiency Syndrome (HIV/AIDS).

The global population is presently estimated to be in the region of 6.7 billion people, of whom approximately 700 million reside in the developed countries of the world (http://www.census.gov/ipc/www/). This demographic is alarming, since it is the developing countries that also have the highest burden of TB disease (WHO Report 2007). South Africa is currently listed as a country with one of the highest TB incidences (WHO Report, 2008), and as such, the situation is indicative of a developing country: - it has a high burden of TB disease, coupled to an increasingly expanding population. Recent statistics released by the Health Systems Trust, a local non-governmental organization (http://www.hst.org.za/healthstats/161/data) cite the national TB incidence as 722.4 per 100 000 population, with the highest TB incidence recorded in the Western Cape and KwaZulu-Natal provinces (in excess of 1000 per 100 000 population). The lowest TB incidences were recorded in Limpopo and Mpumalanga provinces, at 305.2 and 463.4 per 100 000 population respectively.

In South Africa, the treatment success rate has unfortunately been relatively low in the past, reaching only 68% in 2002 (http://www.worldlungfoundation.org/map_SA.php). This indicates that many patients are lost to follow-up health care, and subsequent monitoring for treatment compliance, thereby possibly worsening the local epidemic. Despite these obvious operational difficulties, reasons behind the skewed disease demographics in South Africa as well as other developing countries of the world are still actively debated, given the higher recorded incidences of disease within certain localised areas of these countries.

This debate is furthermore fuelled by the often quoted tenet which holds that only approximately 5-10% of *Mtb*-infected individuals become clinically ill with disease, in the absence of confounding factors, like malnutrition, and/or immunosuppressive conditions such as HIV/AIDS. These confounding factors notwithstanding, this tenet implies that the majority of infected individuals are able to mount some level of "resistance", sufficient to preclude the development of clinically overt disease. On the other hand, *Mycobacterium tuberculosis* strain analysis from different regions of the world has given rise to the idea of strain evolution within (and potentially along with) certain niche populations [3, 4]. Under the precepts of this hypothesis it is envisaged that this host-pathogen interaction could contribute to the selective prevalence of certain *Mtb* strains within these population groups [5-7].

Investigations of the mummified remains of ancient human populations from both Egypt and Peru [8, 9] indicate that TB is an age-old disease. TB palaeomicrobiology, as this branch of the research is known, strives to unravel the evolution of *Mtb* in its co-existence with humankind. An additional impetus was provided by the deciphering of the *Mtb* genome [10], and subsequently, several publications have investigated the global dissemination of *Mtb* strains in different populations [5, 11, 12]. These molecular techniques have assisted in the establishment of modern day TB research practises, which encompasses the development of new TB vaccines, the improvement of the existing drugs and treatment regimens, as well as analysing the factors underlying the pathogenesis and global dissemination of *Mtb*. [13]

Despite this long co-existence with humankind, *Mtb* is still a powerful parasite of its human host, and it can be controlled within populations by the complementary strategies

of early disease detection, (rapid diagnosis) and patient compliance during the full course of treatment. Of these, it is poor patient compliance to treatment that is often highlighted as an important factor contributing to the continuation of the pandemic, and the subsequent emergence of MDR-and XDR-TB [2]. Whilst it is generally accepted that there is a need to improve upon the current practices for disease detection and diagnosis, poor compliance is not fully understood and/or investigated. This is all the more important in the light of the synergy between TB and HIV/AIDS, since an individual afflicted with the dual infections of TB/HIV is required to adhere to a long, onerous therapeutic regimen.

1.2 Isoniazid: An important frontline Antibiotic

Isoniazid, or isonicotinic-acid hydrazine (INH), has been an integral component of all forms of TB therapy since the early 1950s [14]. The continued use of INH is supported by its proven therapeutic efficacy [15], as well as for its role in precluding the development of resistance to co-administered TB drugs in *Mtb* [16]. INH is inexpensive to manufacture, due to its simple chemical structure, and is highly active at easily attainable concentrations that are non-toxic to most individuals [17]. Furthermore, it also exhibits a good overall distribution in a wide range of human tissues, which is very important for counteracting disseminated TB. However, notwithstanding its proven efficacy as an anti-TB drug, humans have the inherent ability to diminish the activity of INH, via an acetylation reaction catalysed by the enzyme arylamine *N*-acetyltransferase 2 (NAT2) [18, 19]. This apparent dichotomy may contribute to the factors confounding the global TB pandemic.

The majority of *Mycobacteria* are sensitive to INH concentrations of greater than 1µg/ml [15]. Most strains of *Mtb* however, are even more sensitive to INH, generally ranging between concentrations of 0.02 µg/ml and 0.06 µg/ml [20]. At present, it is not understood why *Mtb* exhibits such an extreme INH sensitivity in comparison to the other *Mycobacteria*, and numerous attempts have been made to explain the mechanisms for the antimycobacterial action of INH [21-23]. The current consensus proposes that INH is

activated in the mycobacterial cell via an oxidative reaction catalysed by the catalase-peroxidise enzyme, *katG* [24]. Subsequently the activated-INH, which is bactericidal, inhibits the synthesis of mycolic acid, the major cell wall component [25].

In contrast to the above, it has recently been shown that *Mtb* is also capable of inactivating INH by an acetylation reaction catalysed by its own inherent arylamine *N*-acetyltransferase (*NAT*) gene [26, 27]. This gene is highly homologous to the human *NAT2* gene [28, 29], and its presence has also been demonstrated in *Mycobacterium smegmatis* [26] as well as in *Mycobacterium bovis* BCG [30]. The significance of the mycobacterial NAT gene is as yet poorly understood. Nevertheless it appears that the gene is located within an operon that is essential for the survival of this organism within the macrophages of the human host [31].

At the structural level, the human and mycobacterial NAT enzymes both possess a conserved triad of amino acids within the active site pocket of the enzyme, with highly conserved domains situated proximal to this site [27, 32, 33]. Furthermore, initial investigations indicate that prior acetylation of INH by *Mtb*-NAT protects the organism against the activated-INH formed by the *katG* gene product [34].

1.3 The Arylamine *N*-acetyltransferases

1.3.1 The Human Acetylation Phenotype

The first reports describing the metabolic fate of INH in humans were published in 1953 [18], and whilst these studies proclaimed the effectiveness of INH as an anti-TB drug, they also indicated that TB patients differed in their capacity to metabolise INH. During this metabolism a proportion of INH is converted into acetylated-INH, which is ineffective against the tubercle bacillus [19]. Subsequently, a number of studies in healthy individuals, as well as family-based analyses, suggested that the acetylation phenomenon was bimodally distributed in the Caucasian population, indicating that two major alleles are responsible for this acetylation phenotype [28, 35, 36]. In an early analogous study of

Oriental subjects however, the acetylation phenotype was shown to follow a trimodal distribution curve [37-39], indicating that the two alleles could be co-dominant. These early studies served to highlight possible ethnic differences in the acetylation reaction, and facilitated further research into the mechanism of drug metabolism via acetylation [40].

The early investigations employed several xenobiotic compounds to probe the mechanism of acetylation in humans, and phenotypic studies using sulfamethazine (SMZ) [41] procainamide (PRO) [42], and caffeine [43], have confirmed this inter-individual variation in the rate of acetylation. However, similar studies using 4-aminosalicylic acid (4AS) and para-aminobenzoic acid (PABA) as probe substrates observed an invariant monomorphic acetylation phenotype [44]. Thus, human beings have the capability to exhibit both a monomorphic (invariant) as well as a polymorphic (variable) acetylation capacity, depending upon the probe substrate used (Fig. 1.1). These results suggested that separate enzymes, with quite different substrate specificities, could be responsible for the different acetylation phenotypes.

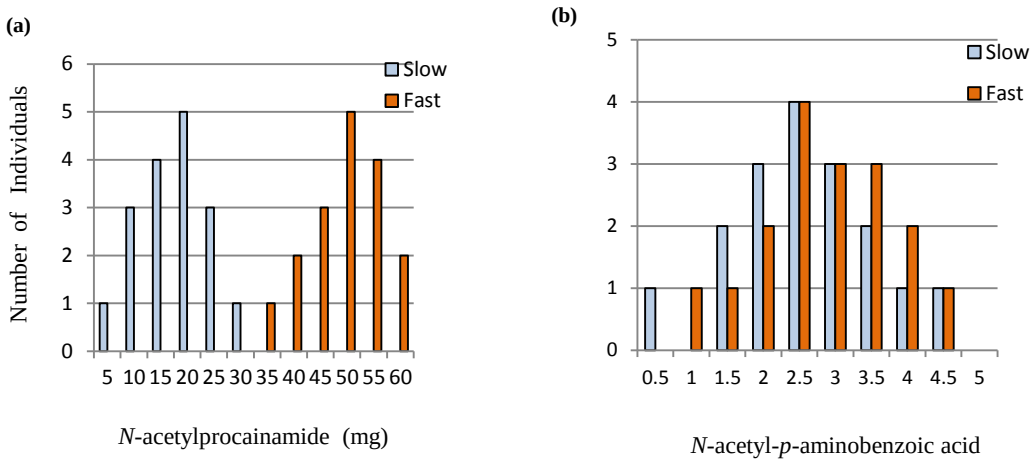


Figure 1.1 Analysis of human acetylation phenotype using procainamide (a) and *para*-Aminobenzoic Acid (b) as probe substrates. The results indicate the polymorphic (a) and monomorphic (b) acetylation frequency distribution. The y-axis represents number of individuals assayed [45].

1.3.2 Monomorphic versus Polymorphic Acetylation

The acetylation reaction has subsequently been shown to be catalysed by the arylamine *N*-acetyltransferase enzymes (NATs), under the Enzyme Commission Classification of E.C. 2.3.1.5 [46]. The NATs are class II drug metabolising enzymes (DME) occurring primarily in the hepatic cell cytosol [47, 48], where they catalyse the *N*-acetylation, *O*-acetylation, as well as the *N*-*O*-transacetylation of xenobiotic compounds possessing a primary aromatic amine or hydrazine substituent group [49-52]. Biochemical analyses indicate that the acetylation reaction proceeds via a two step "ping–pong Bi-Bi" reaction mechanism [53-55], resulting in the transfer of the acetyl moiety from the cofactor acetyl coenzyme A (CoA) to specific reactive groups within the arylamine substrates. The course of the reaction proceeds via the initial formation of a short-lived enzyme-thio-ester intermediate (acetyl-NAT), followed by the subsequent transfer of the acetyl group from the acetylated-NAT to the primary amine nitrogen of the substrate (Fig. 1.2).

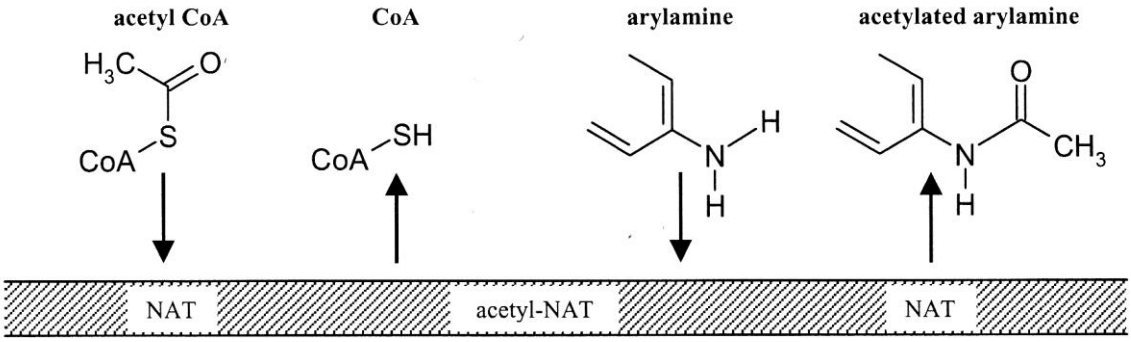


Figure 1.2 The "ping-pong" reaction mechanism for NAT-catalysed acetylation. The acetyl moiety from CoA is initially transferred to a cysteine residue in the active site of the enzyme (acetyl-NAT intermediate), before being transferred to the arylamine substrate (acetylated arylamine) [55-57].

Depending upon the particular arylamine substrate, and whether it undergoes *N*-acetylation, *O*-acetylation or *N*-,*O*-transacetylation, the metabolites formed during this reaction are either inert by-products, which are readily excreted from the body, or are converted into toxic and highly reactive acetoxy and N-acetyl-acetoxy esters [58]. These latter substrates undergo spontaneous decomposition into DNA-or protein-binding nitrenium ions [59], which can initiate carcinogenic and mutagenic changes [58, 60, 61].

These early *in vitro* studies also established that a range of different compounds are metabolised via the acetylation mechanism, including therapeutic agents (dapsone, hydralazine, isoniazid, procainamide, sulfamethazine), important industrial chemicals (2-aminofluorene, benzidine), and environmental toxins such as 4-aminobiphenyl, a constituent of tobacco smoke (Fig. 1.3).

On the basis of the *in vitro* polymorphic acetylation of compounds such as isoniazid and sulphamethazine, individuals can be characterised as either Rapid or Slow acetylators [46, 62], or as Rapid- (RA), Intermediate- (IA), or Slow acetylators (SA) [63-65]. In contrast to these results, acetylation analyses using substrates such as 4-aminosalicylic acid (4AS) and *para*-aminobenzoic acid (PABA), could only demonstrate a relatively small difference in acetylation capacity between individuals, and subsequently these latter substrates are unable to distinguish between the polymorphic categories (of fast, intermediate or slow), as seen with INH [44, 45]. This difference in acetylation capacity is now known to be due to the presence of two separate enzymes, NAT1 and NAT2, which respectively catalysed the selective monomorphic and polymorphic acetylation of substrates seen in man [28, 66-68].

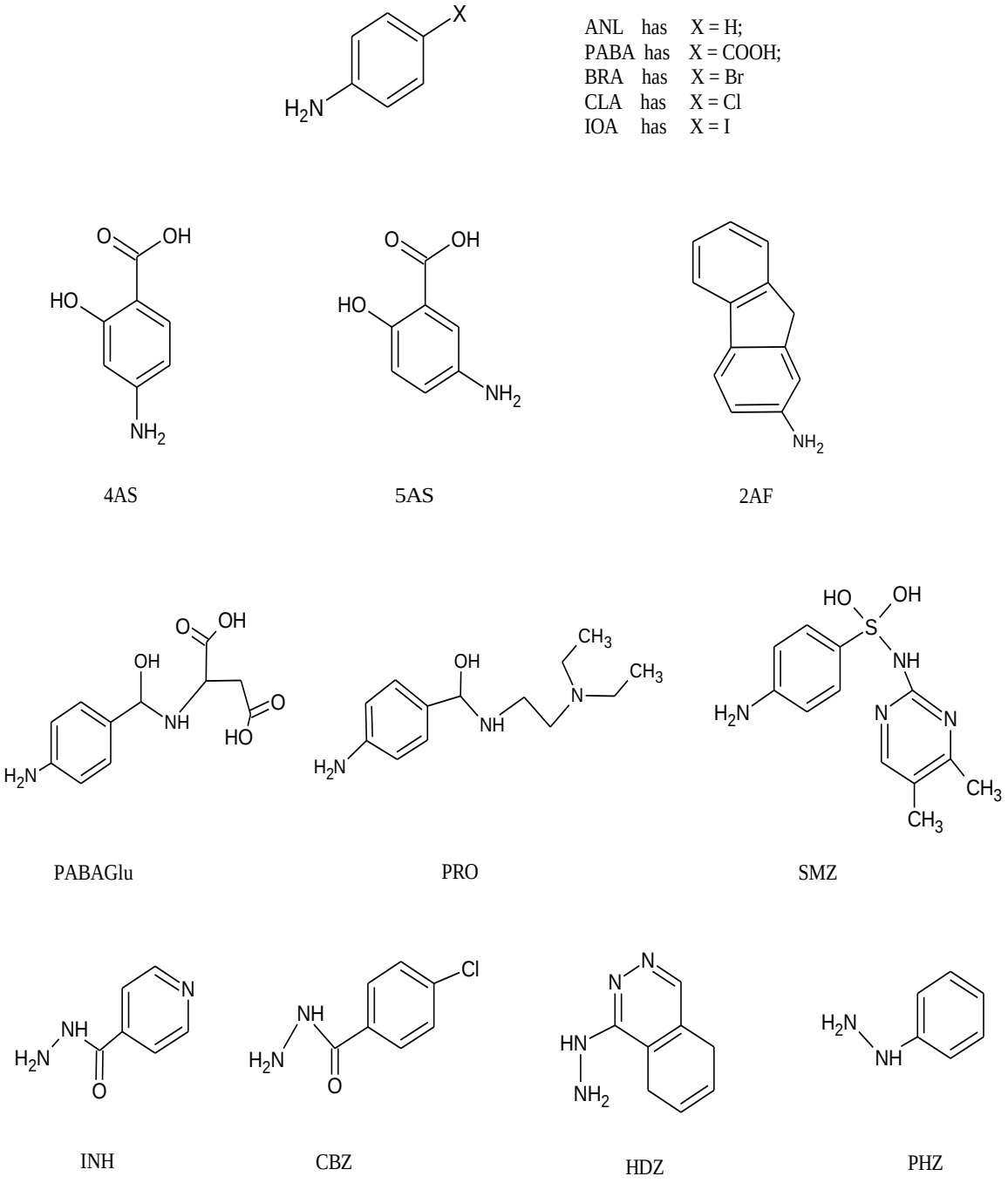


Figure 1.3 Structures of some compounds that undergo NAT-specific acetylation. Compounds are abbreviated as follows:- aniline (ANL), *para*-aminobenzoic acid (PABA), 4-bromoaniline (BRA), 4-chloroaniline (CLA), 4-iodoaniline (IOA), 4-aminosalicylic acid (4AS), 5-aminosalicylic acid (5AS), 2-aminofluorene (2AF), *para*-aminobenzoyl-L-glutamic acid (PABAGlu), procainamide (PRO), sulphamethazine (SMZ), isoniazid (INH), 4-chlorobenozoic hydrazide (CBZ), hydralazine (HDZ) and phenylhydrazine (PHZ).

This early notion of monomorphic versus polymorphic acetylation has since been refuted by several studies, which indicated that NAT1 acetylation in leucocytes [69], liver tissue [70, 71], and whole blood [72], also exhibits an inter-individual variation in acetylation. Several additional studies have confirmed this polymorphic acetylation capacity of the NAT1 enzyme [73-77]. Currently the consensus amongst researchers seems to be that both the NAT1 and NAT2 enzymes are polymorphic [68], due to the presence of several single nucleotide polymorphisms (SNPs) within these genes, that specifically alter the activity of the enzyme. Subsequently, numerous epidemiological studies have sought to elucidate the role of human NAT acetylation in relation to susceptibility to infectious and malignant diseases [78-81]. However, the overall clinical significance of this apparent difference in the human acetylation capability is still unclear.

The acetylation of therapeutic agents such as INH, dapsone (used in the treatment of leprosy, and certain forms of malaria [82]), and procainamide (used in the therapy of cardiac arrhythmias [42]), impacts upon the therapeutic treatment of the patient. Subsequently it has been established that some patients are prone to the development of adverse drug reactions (ADR) and chemically induced toxicity caused by their inherent metabolism of these drugs. This can require that the drug(s) be withdrawn from treatment, which could imply subsequent less effective treatment. Secondly, in the case of TB and leprosy, such partial exposure to the drugs, could facilitate the development of drug resistance in the disease-causing organism. Hence it is important to develop an understanding of the metabolism of these compounds with respect to acetylation, because the drugs in question feature prominently in the treatment regimens for these diseases.

1.3.3 The Human *NAT* Genes

Three *NAT* genes have been detected and characterised in humans, namely, *NAT1*, *NAT2*, and *NATP* [83]. These genes occur on chromosome 8, in the region of 8p22 [84-86]. All three genes occur relatively close together, within a 360 kilobase pair (kb) domain (Fig. 1.4), with the chromosome-specific marker D8S261 situated telomeric to the

NAT1 gene [86]. Functional NAT enzymes are derived from single open reading frames (ORF) of 870 nucleotides in length [70]. These protein coding domains share a nucleotide sequence identity of 87%, and code for proteins of 34 kiloDaltons (kDa) containing 290 amino acids which are 81% identical [83].

Early studies using hepatic tissue indicated that the messenger RNA (mRNA) transcripts for *NAT1* have an exact co-linearity with a single exon locus, implying that the protein transcript is entirely derived from this exon. On the other hand, the *NAT2* mRNA transcript consists of sequences from the primary protein-coding exon, as well as sequences from a small non-coding exon, located 8.65 kb upstream of it [87] (Fig. 1.4). In contrast, the *NATP* locus contains numerous stop codons, and neither immunoreactive protein nor acetylation activity could be detected in *NATP in vitro* transfection experiments [83]. Consequently, *NATP* is believed to be a pseudo-gene.

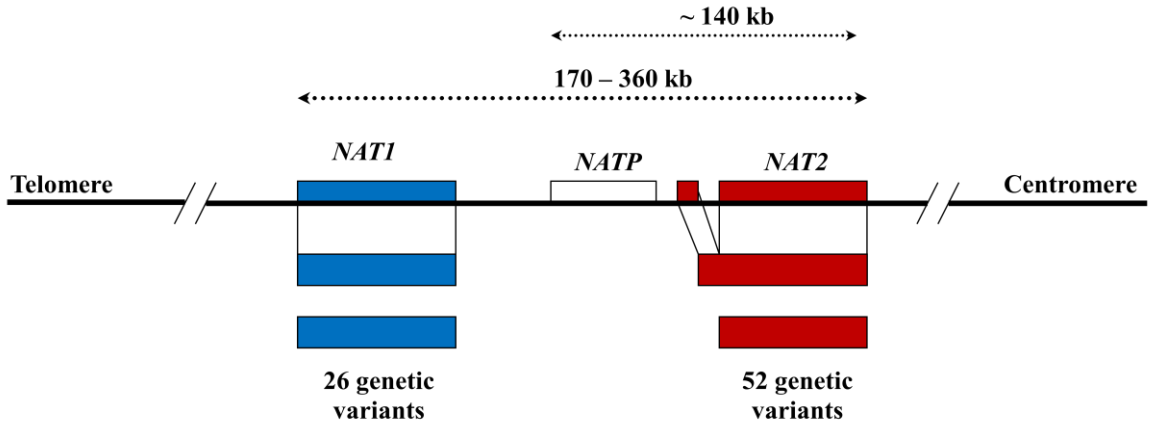


Figure 1.4 The organisation of *NAT* gene loci on Chromosome 8 [86].

Both *in vivo* and *in vitro* studies have established that the expression of *NAT1* occurs independently of *NAT2* [88], which indicates that these genes are under the control of separate promoters. There is however, a paucity of published information on the mechanism(s) controlling the expression of the *NAT* genes. Nevertheless, it has been established that NAT1 is expressed in diverse tissues, such as liver [70], leukocytes [69], bladder [89], erythrocytes [90], placenta [91], foetal tissue [92], and intestinal epithelium

[77]. The *NAT2* gene also exhibits a diverse tissue expression profile, but generally at a relatively lower rate of transcription compared to NAT1, in most tissues. A higher level of NAT2 expression has however been found in the liver, colon and small intestine [93, 94]. This relatively high level of NAT2 expression in these tissues, traditionally regarded as sites of initial drug metabolism, may support the hypothesis that NAT2 is primarily involved in the metabolism of various drugs. On the other hand, the wide-spread expression of NAT1 in human tissue suggests that NAT1 may fulfil an inherently endogenous role within the body.

The significant structural similarity between NAT1 and NAT2 [95] suggests that these enzymes could also exhibit an overlapping substrate selectivity, and this has been shown using various probe substrates [96]. A recent publication [97], investigating the substrate specificity of the human NAT enzymes has provided support for the overlapping substrate hypothesis, by observing that, whilst many of the probe substrates tested exhibit a clear NAT1- or NAT2-acetylation specificity, some of the compounds were equally acetylated by both enzymes (Fig. 1.5). Despite this overlapping substrate specificity, it is still not known whether NAT1 and NAT2 act co-operatively and/or sequentially within the same or disparate acetylation pathways.

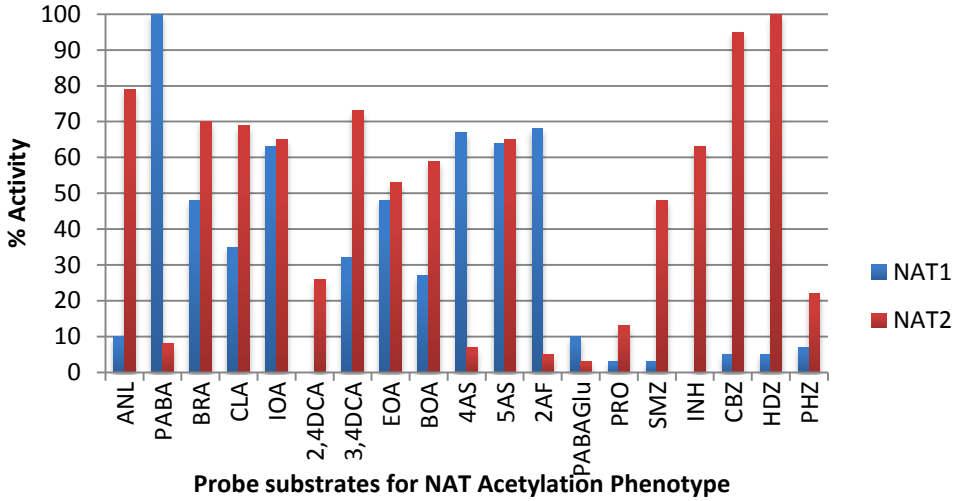


Figure 1.5 Substrates screened for NAT1 and NAT2-specific acetylation via acetyl-coA hydrolysis. Abbreviations for the different compounds are:- aniline (ANL), *para*-aminobenzoic acid (PABA), 4-bromoaniline (BRA), 4-chloroanaline (CLA), 4-iodoaniline (IOA), 2,4-dichloroaniline (2,4DCA), 3,4-dichloroaniline (3,4DCA), 4-ethoxyaniline (EOA), 4-butoxyaniline (BOA), 4-aminosalicylic acid (4AS), 5-aminosalicylic acid (5AS), 2-aminofluorene (2AF), *para*-aminobenzoyl-L-Glutamic acid (PABAGlu), procainamide (PRO), Sulphamethazine (SMZ), isoniazid (INH), 4-chlorobenzoic hydrazide (CBZ), Hydralazine (HDZ), and phenylhydrazine (PHZ). Adapted from [97].

1.3.4 Structure and Functional Studies of NAT Proteins

The NAT proteins constitute a highly conserved family of enzymes that are well represented amongst both prokaryotic and eukaryotic organisms, including humans [93], chickens [98], rabbits [99, 100], rats [101], cats [102] and various bacterial species [26, 103]. Interestingly, dogs and other canids do not exhibit any *N*-acetyltransferase activity [104]. This substantial body of early knowledge has facilitated our understanding of the NAT acetylation mechanism. Across species comparisons of the NATs have identified a number of highly conserved amino acids, as well as functionally important structural motifs (Fig. 1.6). Due to this conserved sequence identity amongst the NATs, a number of model organisms have been developed to study human acetylation. In this context, studies of bacterial NAT [52] have shown that these enzymes contain a catalytic triad of amino

acids (Cys[69]–His[107]–Asp[122]) as the primary feature of the enzyme's active site pocket [32, 33]. In human NAT, this structurally and functionally analogous feature is depicted by the Cys[68]-His[107]-Asp[122] triad [105, 106].

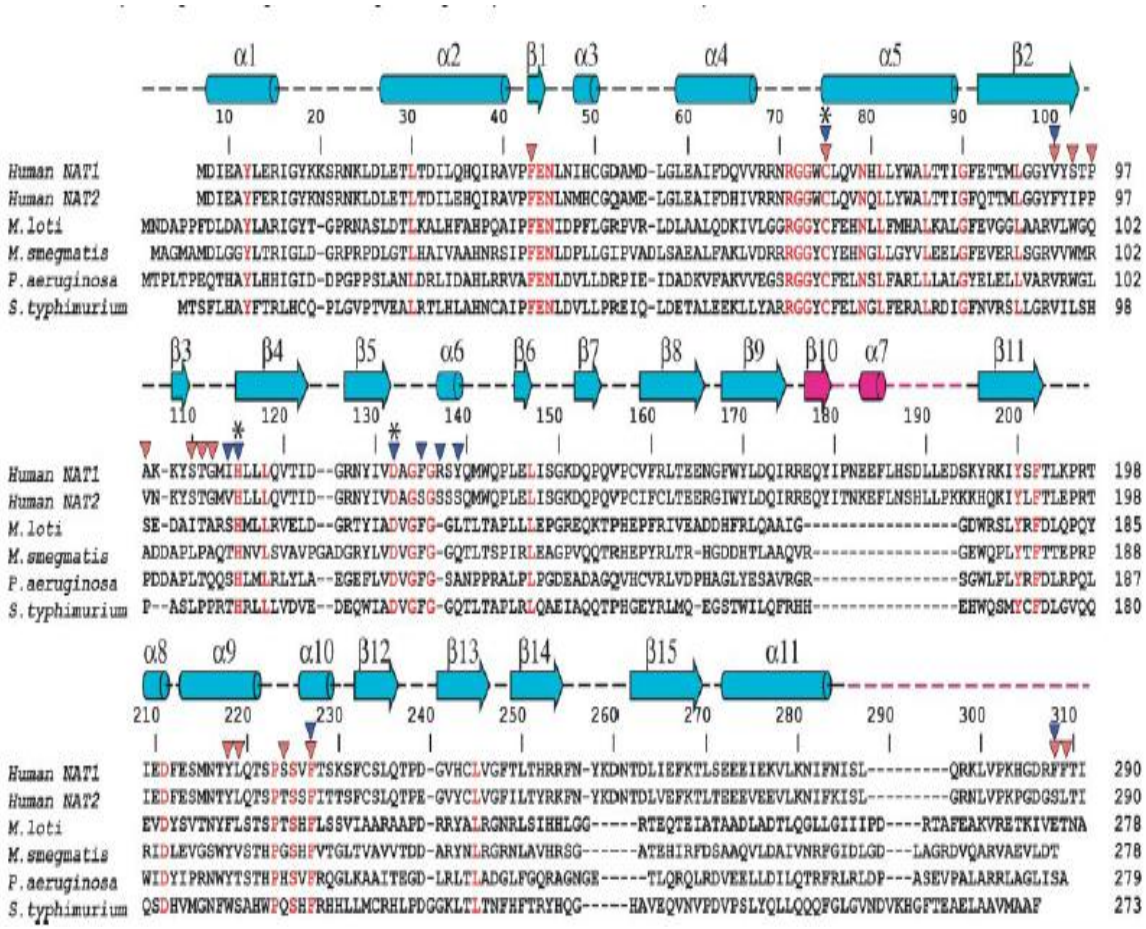


Figure 1.6 A sequence alignment of NAT proteins from several different species to show the overall structural similarity between them - mostly α-helical (cylindrical) and β- sheet (arrow bar) conformations. The magenta colour (for example in the β10 and α7 domains) indicate regions specific to human NATs. Highly conserved amino acids are indicated in red; Acetyl CoA-binding amino acids are indicated as red triangles, whilst substrate-binding amino acids are indicated with blue triangles (taken from Wu *et al* [107]).

Despite the similarities in overall structure, crystallographic experimentation and computational analyses have revealed the presence of two additional structural features in the human enzymes:- a 17-residue inserted domain (amino acids 167-183), and a domain at the carboxy terminus which assumes a random coil conformation [107, 108]. These domains are not present in other Eukaryotic species, nor have they been found in any of the bacterial NATs thus far studied (Fig.1.7). It is currently speculated that these domains may feature prominently in the catalytic activity of the human enzymes, since the 17-residue inserted domain is orientated in such a way as to interact with several of the β–sheet and α–helical domains in the native enzyme structure, thereby contributing to the overall stability of the human proteins (Fig.1.7).

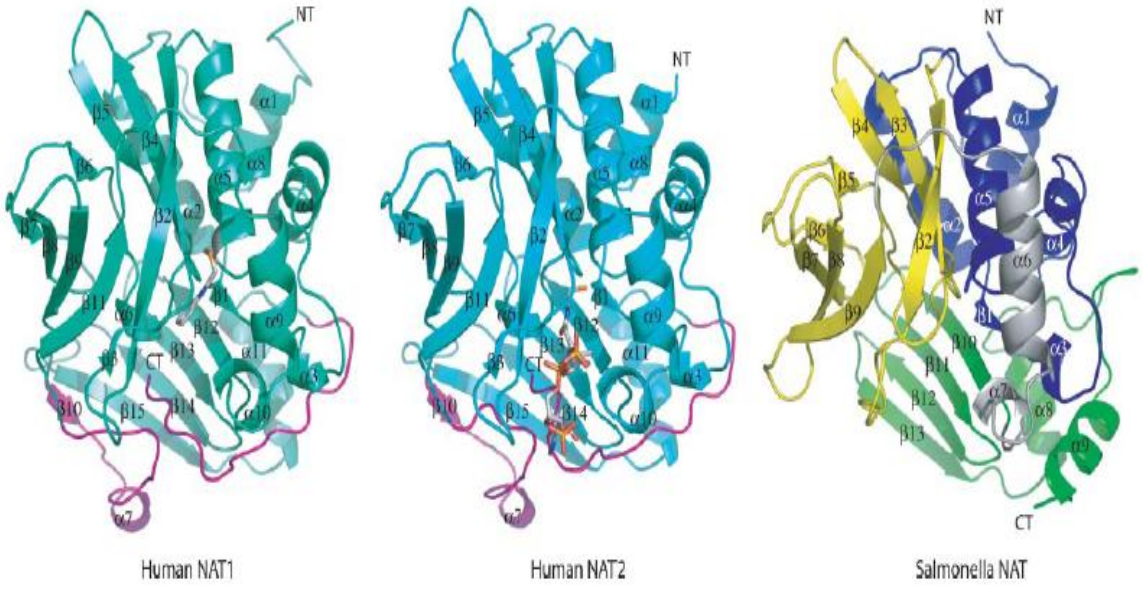


Figure 1.7 Ribbon diagrams of the overall structure for human and bacterial NATs, showing the domains of difference (magenta) prevalent in human NATs (ref. [107]).

The carboxy-terminal region assumes an extended structural conformation, and stretches across the opening of the catalytic site pocket, effectively shielding the pocket and restricting access to this site. This orientation furthermore promotes the initial interaction between human NAT and acetyl-CoA, during the formation of the characteristic acetylated enzyme-substrate intermediate [109]. Furthermore, this carboxy-terminal domain, specifically from amino acid 274 onward, is oriented into the interior of the molecule, and is thus effectively shielded from interacting with the aqueous outer

environment [109]. In contrast, the carboxy terminal domain in Prokaryotes occurs on the periphery of the enzyme molecule, away from the catalytic site pocket [32]. Despite the overall structural similarities between the Eukaryotic and Prokaryotic NAT enzymes, the significance of these latter structural differences have yet to be determined.

Studies investigating the substrate binding selectivity of NAT1 and NAT2 observed that the carboxy terminal domain of the protein, along with residues 124 to 129, facilitates the interactions of the substrate with the active site of the enzymes [110, 111]. According to Wu *et al* [107], the amino acids at position 93, 125, 127 and 129 feature prominently in the substrate selectively of the human enzymes by affecting the interactions within the catalytic site pocket. In the case of NAT2, the amino acid serine occurs at positions 125, 127 and 129. In contrast to this, the NAT1 protein has phenylalanine, arginine and tyrosine at these respective positions [107]. These authors also show that the bulky side chain groups of arginine-127 and tyrosine-129 protrude into the NAT1 active site pocket, effectively reducing the substrate binding volume by 40% with respect to the NAT2 substrate binding pocket. Consequently, the NAT1 substrates are relatively small, compared to NAT2-specific substrates (see Fig.1.3, PABA vs SMZ). Furthermore, the amino acid valine occurs at position 93 in NAT1, whereas phenylalanine occurs at position 93 in NAT2. This difference affects the binding capacity within the substrate binding pocket of NAT2 specifically, making it more conducive for the interaction with hydrophobic substrates, due to the flat planar surface of the benzene ring structure in the side chain group of phenylalanine.

The substrate binding specificity of the human NATs also appears to be largely influenced by the identity of the amino acid at position 125. In a series of substitution experiments conducted by Goodfellow *et al* [111], the substitution of phenylalanine-125 with serine in the NAT1 enzyme, abrogated the usual NAT1-specific interaction with PABA, and resulted in a NAT1 enzyme that specifically interacted with SMZ, a NAT2-specific substrate.

Besides the significance of particular amino acid residues within these enzymes, the importance of the catalytic site triad (Cys-His-Asp) cannot be over emphasised.

Homology modelling studies have shown that these three amino acids are conserved in all known NAT enzymes, and that the cysteine residue in particular, is crucial for the acetylation reaction [52, 105]. As indicated earlier, this cysteine residue is the site of initial acetylation of the enzyme with acetyl CoA, prior to the acetylation of the substrate. The significance of this aspect of the reaction is borne out by the observation that the acetylated NAT moiety is more stable than an un-acetylated enzyme. Moreover, prior acetylation of the substrate rather than the enzyme, leads to the rapid degradation of the enzyme via the ubiquitin-proteasome pathway [112]. Hence, whilst the substrate selectivity may be prescribed by the volume of the catalytic pocket, as well as by interactions with particular amino acid residues, the overall rate of the acetylation reaction may be specified by the initial interaction between acetyl CoA and the enzyme moiety.

1.3.5 NAT Expression Studies

The human *NAT* genes, despite their similarities, appear to have very distinct functional roles, based upon their observed expression profiles in different tissues, and during the developmental stages of the organism. These expression profiles demonstrate apparent tissue selectivity.

Recently, a NAT knockout mouse model [113] was used to broadly investigate NAT expression, as well as deduce the role of these enzymes in Eukaryotes during embryonic development. The murine and human NAT proteins are highly analogous [114], and map to syntenic regions on chromosome 8 in both species [115-117]. These factors make the mouse a particularly useful model system for studying acetylation, with murine Nat1 exhibiting a substrate specificity equivalent to human *NAT2*, whilst murine Nat2 is equivalent to human *NAT1*. Loehle *et al* [118], in using this Nat2 knockout mouse model, found no significant difference in relative expression levels (mRNA) for both Nat genes, assayed via quantitative polymerase chain reaction analysis, in the murine hepatic and extrahepatic tissue compared to the wild type strain. This result was consistent with similar studies in human tissues [116, 119]. Secondly, using specific probe substrates to assess the acetylation activity – INH for Nat1 and PABA for Nat2 – these authors were able to show that the reduction in Nat2 catalytic activity was caused by the consequential

reduction in Nat2 protein. Interestingly, Nat1 transcription (mRNA) and catalytic activity (protein) did not reflect a compensatory adjustment as a result of the loss of Nat2 in the knockout strain. Taken together, these results indicate that the NAT acetylation reaction fulfils important and similar endogenous roles in the xenobiotic biotransformation mechanism in Eukaryotic tissue.

1.3.5.1 *NAT1* Gene Expression

The *in silico* analysis of expressed sequence tags (ESTs) for the human contig NT_030737 (http://www.ncbi.nlm.nih.gov/nuccore/51466871) has indentified several putative splice variants for the *NAT1* gene. Subsequently, various *NAT1* mRNA transcripts, differing only in the 5'-untranslated region (5'-UTR), have been reported [120-123]. On the whole, the 5'-UTR contains 8 separate exons, giving rise to multiple putative mRNA transcripts, consisting of various combinations of these exons conjoined to the single protein coding exon, exon 9 (Fig. 1.8).

Studies of the human *NAT1* gene locus have identified at least two promoter domains, one of which is located approximately 12 kb upstream of exon 9, called NATb [120, 123], and another, NATa, located 51.5 kb upstream of this protein coding exon [121]. The most proximal promoter domain, NATb, occurs just upstream of exon 4. It contains a putative binding site for the Activator Protein 1 (AP-1) transcription factor, flanked by a –TCATT repeat motif. In addition, a functional transcription factor binding site for Stimulatory protein 1 (Sp1) has also been identified within this NAT1 transcriptional domain. Since the Sp1 transcription factor is ubiquitously found in association with house-keeping genes [124], it is believed that the NATb promoter region drives the constitutive expression of NAT1 [123]. The second promoter, NATa, appears to facilitate a higher level of *NAT1* expression in lung, liver, kidney and tracheal tissues, and hence, may function at a more tissue-specific level. Thus the expression of *NAT1* proceeds from two separate promoters, which appear to exhibit differential activity in different tissues, supporting the idea that the expression of *NAT1* is both inducible [125], and/or constitutive [123].

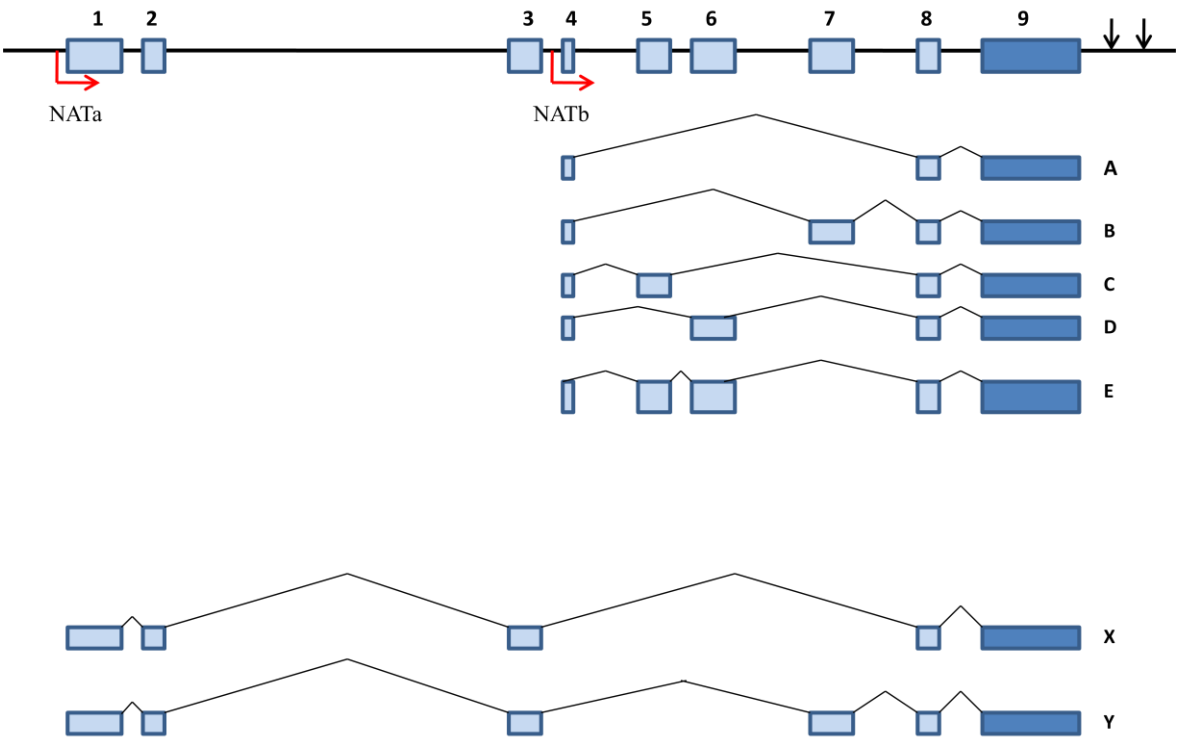


Figure 1.8 Diagram depicting the genomic arrangement for the human *NAT1* locus, adapted from Minchin *et al* [126]. Several mRNA splice variants (A to E, X and Y), arising from the separate promoters NATa and NATb (red arrows), have been described. The *NAT1* protein coding sequence is derived exclusively from exon 9 [114, 126]. Putative poly-adenylation sites are also indicated (black arrows).

1.3.5.2 *NAT2* Gene Expression

In the case of the *NAT2* gene, a different restricted expression profile is observed, with the highest level of expression occurring in the liver, colon and small intestine [127]. In their study of full length mRNA splice variants from 29 different tissue types, Husain *et al* [127] observed baseline expression levels only in the remainder of the human tissues analysed. They also observed that the majority of these mRNA transcripts were expressed from a single promoter, located 8.7 kb upstream of the *NAT2* open reading frame (ORF) (Fig. 1.9) [87]. In addition, the *NAT2* transcription start site (TSS) was located just upstream of the non-coding *NAT2* exon, within a 100 nucleotide domain [120].

In contrast to these observations, Boukouvala and Sim [116] found that *NAT2* transcripts may also be initiated from a second promoter, located immediately 5' to the *NAT2* ORF.

The general importance of the latter promoter has yet to be established, since its prominence has thus far only been demonstrated in *NAT2* transcripts from liver tissue, and not in extra-hepatic tissue [127]. Interestingly, the very low level of *NAT2* expression observed in the other tissues analysed by Husain *et al* [127], may well indicate that *NAT2* expression could be under the influence of a relatively weak promoter in these tissues, thereby lending support to the findings of Boukouvala and Sim [116]. The different expression profiles that have been observed may require the additional intervention of tissue-specific transcription factors.

A comparison of the 8.7 kb *NAT2* promoter domain with the corresponding region in the paralogous *NAT1* gene has shown that they share an 85% sequence homology. Despite this significant similarity, the *NAT2* transcription start site (TSS) domain does not contain the binding site for the Sp1 transcription factor [127]. These observations indicate that the expression of the *NAT2* gene is inducible rather than constitutive, thereby supporting the findings of differential *NAT2* expression in diverse human tissue.

Generally speaking there is a limited amount of information available on the mechanism(s) affecting the regulation of *NAT2* gene expression. However *in silico* analysis of the reported genome sequences have identified several other putative transcription factor binding sites in the 5'-UTR of *NAT2*, as well as tandem polyadenylation sites in the 3'-UTR of the gene [114, 116, 128]. The functional importance of these binding sites has yet to be analysed, and therefore the mechanistic role of the 5'-UTR in *NAT2* gene expression is as yet undefined. In a recent study of *NAT2* polymorphisms in the Indonesian population, several novel SNPs were reported in the 5'-UTR of the gene [129]. However similar observations in other population groups have not been reported, and hence the importance of these SNPs in the regulation of *NAT2* expression has yet to be established.

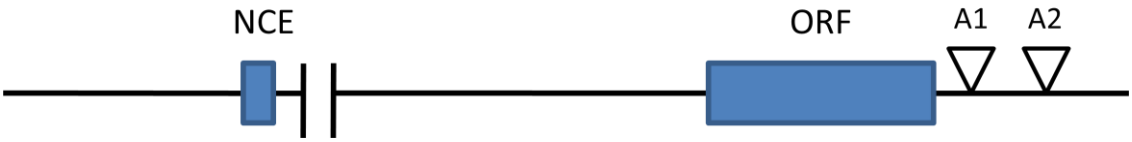


Figure 1.9 The genomic arrangement for the human *NAT2* gene locus. A single non-coding exon (NCE) is located approximately 8.7kb upstream of the open reading frame (ORF) of the protein coding exon. Open triangles indicate two putative poly-adenylation signalling sites (A1, A2).

NAT2 mRNA is differently polyadenylated in the 3'-UTR in various vertebrate species [100, 130-132], thereby affirming the presence of the tandem polyadenylation sites. In the case of the paralogous *NAT1* gene, the A2 adenylation site (Fig.1.9) appears to be the preferred site for mRNA polyadenylation [116], since *in silico* analyses of *NAT1* ESTs have indicated that the polyA tail originates at the polyA2 site. The preferential utilisation of this polyA site could imply that the subsequently formed mRNAs may be more stable. Whilst the mechanism for the post-transcriptional regulation of *NAT2* has yet to be elucidated, it is likely that a similar mechanism ensures the differential polyadenylation of the *NAT2* mRNAs.

Alternative splicing of mRNA transcripts is a general property of the cell's translational processes, and is believed to be a feature of cell specific gene expression and regulation [133, 134]. As such, the alternative splicing of the *NAT* mRNA transcripts is not peculiar to the human species, and has also been demonstrated in hamsters [130], rabbits [99], and rats [131]. However, the aspect of alternative splicing in intronless genes may be worthy of further investigation, given that less than 5% of all genes in the higher order Eukaryotes are believed to be intronless [135].

1.3.6 Human NAT Acetylation Polymorphisms

Many investigations have attempted to establish whether there is a causative association between various diseases and the individual's acetylation capacity with respect to one or both of these enzymes (NAT1 and NAT2). For example, NAT acetylation has been investigated in a variety of different cancers, such as bladder [136], breast [137], colon [138], and prostate [139]. Currently the basis for the acetylation polymorphism is ascribed to numerous activity-altering SNPs that have been identified within the NAT genes. These SNPs have been widely investigated in various studies for their contributory effects towards the differential acetylation of drugs and xenobiotics [28, 78, 128, 140].

1.3.6.1 The NAT1 Acetylation Polymorphism

At least 25 variant alleles have been characterised in the *NAT1* gene. By convention, the reference allele, *NAT1*4*, confers the "wild-type" acetylation activity observed for the enzyme (http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/). A total of 32 different nucleotide changes distinguish these 25 vaiant alleles from *NAT1*4* (Table 1.1). Many of these changes occur as individual SNPs, defining separate alleles (for example as in *NAT1*3* and *1*17*), or occur as combination-SNPs, characteristic of allele clusters or haplotypes (*NAT1*11A, 11B, 11C, and 1*18A, 18B*).

Eighteen different SNPs are localised to the protein coding region of the gene (exon 9, Fig. 1.8), whilst the remaining 14 nucleotide changes occur within the UTRs. Extant information regarding the effect of these polymorphisms on the NAT1 phenotype is, at best incomplete, and the acetylation capacity for at least 9 of these variants namely, *NAT1*5, 1*11C, 1*16, 1*18A, 1*18B, 1*26A, 1*26B, 1*28* and *1*29*, has still to be determined.

Table 1.1. Single nucleotide polymorphisms associated with various *NAT1* alleles.

Nucleotide change(s)	Amino Acid Change	Defining *NAT1* Allele(s)
5'UTR		
-344C>T	None	*1*11A*, *1*11B*, and *1*11C*
-40A>T	None	*1*11A*, *1*11B*, and *1*11C*
Exonic Region		
21T>G	Synonymous (L7L)	*1*27*
97C>T	R33Stop	*1*19*
190C>T	R64W	*1*17*
350G>C, 351G>C	R117T	*1*5*
402T>C	Synonymous (P134P)	*1*20*
445G>A	V149I	*1*11A* and *1*11B*
459G>A	Synonymous (T153T)	*1*11A*, *1*11B*, and *1*11C*
497-499G>C	R166T; E167Q	*1*5*
559C>T	R187Stop	*1*15*
560G>A	R187Q	*1*14A* and *1*14B*
613A>G	M205V	*1*21*
640T>G	S214A	*1*11A*, *1*11B*, and *1*11C*
752A>T	D251V	*1*22*
777T>C	Synonymous (S259S)	*1*23* and *1*27*
781G>A	E261K	*1*24*
787A>G	I263V	*1*25*
3'UTR		
884A>G	None	*1*5*
Δ976	None	*1*5*
Δ1025	None	*1*29*
Δ3 between 1065-	None	*1*18A*
Δ3 between 1065-	None	*1*18B*
(AAA) after 1091	None	*1*16*
(TAA) insert between	None	*1*26A* and *1*26B*
(TAATAA) deletion	None	*1*28*
Δ9 between 1065-1088T>A	None	*1*11A*, *1*11B*, and *1*11C* *1*10*, *1*14A*, *1*18A*, *1*29*
1095C>A	None	*1*3*, *1*10*, *1*11A*, *1*11C*, *1*14A*, *1*16*, *1*18A*, *1*26A*, and *1*29*
Δ1105	None	*1*5*

See List of Abbreviations for explanation of single letter Amino acid abbreviations.
Table adapted from NAT website

A limited number of alleles have however been investigated, of which, only a relatively small subset (most notably, *NAT1*3*, *NAT1*4*, *NAT1*10*, *NAT1*11*, and *NAT1*14*), have been shown to exhibit differences in phenotype, in mostly Caucasian cohort studies [141-144] (see Table 1.2 below). Despite this knowledge from early *in vitro* acetylation studies, there is still a limited understanding of the *NAT1* genotype-phenotype inter-relationships. Thus, acetylation studies based on strictly phenotypic analyses may not be able to characterise a particular *NAT1* variant locus. Therefore it is important to develop improved and more robust experimental methodologies to define the SNP frequencies within populations, in order to accurately define the acetylation phenotype conferred by these SNPs.

1.3.6.1.2. Functional Impact of NAT1 Polymorphisms

The acetylation phenotypes for various *NAT1* haplotypes are depicted in Table 1.2. It is interesting that the alleles with an acetylation phenotype "greater than" *NAT1*4*, show polymorphisms that occur within the UTRs of the gene. Moreover, the *NAT1*10* allele encodes a greater acetylation capacity than *NAT1*3* [116], and it would appear that this difference may be solely due to the additional SNP at position 1088. The significance of the 1088T>A polymorphism may be further extrapolated, based upon the observation that the 1095C>A SNP on its own does not appear to substantially affect the acetylation capacity, since the *NAT1*3* phenotype is observed to be equivalent to that of the reference allele, *NAT1*4*. However, these alleles have not been studied in diverse populations, and the contribution of these SNPs (and alleles) to the overall acetylation phenotype is unknown. Hence, further investigations are necessary to substantiate these results obtained from Caucasian cohort studies [141-144]. In general, the role postulated for NAT1 in the metabolism of xenobiotic compounds has been developed from early studies using validated animal models (reviewed by Sim *et al*, 2008 [128]), and subsequently many investigations have attempted to define a causative association between *NAT1* acetylation polymorphisms and malignant disease susceptibility [145, 146].

Table 1.2 Deduced acetylation phenotypes for polymorphic NAT1 Haplotypes.

Acetylation Phenotype Reference	NAT Allele (Haplotype)	Nucleotide Change	Deduced Amino Acid Change
	NAT1*4	**Reference**	**None (Reference Sequence)**
Equivalent to *1*4*	*NAT1*3*	1095C>A	
	*NAT1*20*	402T>C	P134P‡
	*NAT1*21*	613A>G	M205V
	*NAT1*23*	777T>C	S259S‡
	*NAT1*24*	781G>A	E261K
	*NAT1*25*	787A>G	I263V
	*NAT1*27*	21T>G; 777T>C	L7L‡; S259S‡
Greater than *1*4*	*NAT1*10*	1088T>A; 1095C>A	
	*NAT1*11A*	-344C>T; -40A>T 445G>A 459G>A 640T>G Δ9 between 1065-1090 1095C>A	 V145I T153T‡ S214A
	*NAT1*11B*	-344C>T; -40A>T 445G>A 459G>A 640T>G Δ9 between 1095-1090	 V145I T153T‡ S214A
Lower than *1*4*	*NAT1*14A*	560G>A 1088T>A; 1095C>A	R187Q
	*NAT1*14B*	560G>A	R187Q
	*NAT1*15*	559C>T	R187Stop
	*NAT1*17*	190C>T	R64W
	*NAT1*19*	97C>T	R33Stop
	*NAT1*22*	752A>T	D251V

‡ = Synonymous amino acid change conferred by SNP. Table adapted from the official NAT website.

These latter studies have reported some association, yet the overall results have been tenuous, due mainly to the low resolving power inherent in genetic studies employing a small number of participants, and the difficulty in replicating the results in other cohorts. Furthermore, only some of the *NAT1* alleles have thus far been investigated in a limited number of population cohorts [141-144]. Nevertheless, a seemingly regular association

has been demonstrated between several types of cancers and the *NAT1*10* allele [76, 147-149]. These latter results are of particular interest since the SNPs defining *NAT1*10* all occur in the 3'-UTR of the gene, and by popular convention are not considered to have any effect on the activity of the enzyme [150, 151].

Recent studies employing micro-array technologies have shown an association between an increased NAT1 expression and oestrogen receptor positive breast cancer cells, which would indicate the involvement of a hormone response element, presumably present within the promoter domain of this gene [152, 153]. Thus the study of the polymorphic *NAT1* alleles will continue to contribute valuable insights into the pharmacogenetics of acetylation.

1.3.6.1.3 Ethnic distribution of *NAT1* polymorphisms

Several studies have investigated the frequency distribution for the various *NAT1* alleles, however the existing information is not as comprehensive as for the *NAT2* gene. This imposes a major limitation upon comparative studies investigating the contributory effects of the acetylator polymorphisms in different populations. Furthermore, many of the earlier studies utilised a PCR-RFLP analytical protocol exclusively, which facilitated the investigation of a targeted subset of *NAT1* alleles only, mainly, *1*4*, *1*3*, *1*10* and *1*11* [76, 154-156]. In other instances, the complete *NAT1* sequence domain was investigated via genetic sequencing [157-159], which indicated that some *NAT1* alleles are rare (occurring at a frequency of 1% or less) or absent in these extant populations. Table 1.3 summarises the *NAT1* allele frequencies in different populations.

Despite these limitations, both in the alleles investigated, and in the limited number of ethnicities analysed, it is apparent that the profile of *NAT1* alleles found in Caucasians compared to Asians is sufficiently diverse to account for the observed differences in NAT1 acetylation status for these groups [75, 158]. In this respect, the ethnic African populations have to a large extent been neglected in these analyses, and therefore it is difficult to reach a consensus on the status of the NAT1 acetylation within the latter groups. It is however extremely important to address this limitation since this group

appears to exhibit allele frequencies that are similar to both Caucasians and Asians, as indicated in Table 1.3.

Table 1.3. The distribution of *NAT1* alleles as reported for various ethnic groups.

Ethnic Group	No. Of Alleles tested (n)	Allele frequency expressed as percentage (%)						Reference
		1*4	1*3	1*10	1*11	1*14	Other[Δ]	
Caucasians								
British	224	73.7	3.6	18.7	3.1	0.9	0	[155]
French	540	74.4	2.8	17.8	0	3.7	1.3	[144]
German	628	70.9	3	20.1	3.3	2.2	0.5	[76]
African								
Lebanese[¥]	84	56	3.6	10.7	0	23.8	n.d.	[156]
S.A. Black	202	48.5	1	50.5	0	0	0	[155]
Asian								
Indian	280	51	30	17	2	n.d.	n.d.	[154]
Malay	244	30	29	39	2	n.d.	n.d.	[154]
Thai	466	50.4	3.4	43.8	2.4	0	n.d.	[157]
Han Chinese	280	49.6	8.2	40	2.2	n.d.	n.d.	[158]
Taiwanese	342	54	3.2	42	6	n.d.	n.d.	[160]
Japanese	98	56.2	1.0	40.6	0	0	5.7[†]	[161]

n.d. = Alleles not determined by the experimental assay;
[Δ] = combined frequencies of alleles, such as, *1*15*, *1*17*, etc. usually <1%.
[¥] Cohort consisted of second generation Lebanese individuals living in the United States of America at the time of the study
[†] Figure represents the sum of frequencies specifically for the *NAT1*18* alleles only; no other NAT1 alleles were found in this study.

The importance of elucidating how *NAT1* genotypes affect the acetylation phenotype is also apparent from contrasting results obtained by separate phenotypic studies [75, 110, 141, 158]. Studies by Hughes *et al* [75] and Payton & Sim [110] found the acetylation activity of the *NAT1*10* allele to be similar to that of the *1*4* allele, using PAS and PABA as activity probes respectively. On the other hand, Bell *et al* [141] and Zhangwei *et al* [158], using PABA to probe for NAT1 acetylation, observed an elevated acetylation activity for the *NAT1*10* allele in homozygous as well as heterozygous genotypes. This apparent discrepancy could be ascribed to differences in study design, in that Hughes *et al* [75] and Payton & Sim [110] investigated NAT1 acetylation in Caucasian individuals,

whilst Zhangwei *et al* [158] analysed Chinese individuals, and Bell *et al* [141] conducted an *in vitro* study of bladder and colon tissue samples. Interestingly, the mean NAT1 activity in bladder tissue was 8 times lower than in the colon samples, yet a higher NAT1*10 activity was observed in bladder samples relative to colon tissue [141]. Taken together, these results accentuate the importance of elucidating the *NAT1* genotype-phenotype inter-relationship with further studies.

1.3.6.2 NAT2 Polymorphisms

The classic "isoniazid acetylation polymorphism" has been unequivocally linked to the human *NAT2* genetic locus [28, 36, 46, 50]. Historically, the acetylation reaction has been more frequently studied and characterised with respect to the *NAT2* gene and its well-defined polymorphisms [37, 162-167]. Several of these polymorphisms encode NAT2 proteins exhibiting a reduced enzyme activity [71, 83, 168], giving rise to the Slow acetylator (SA) phenotype. These particular SNPs, namely, 191G>A, 341T>C, 590G>A, and 857G>A, are characterised by complete genotype-phenotype concordance, and have become the defining polymorphisms for the *2*14*, *2*5*, *2*6*, and *2*7* allele clusters (haplotypes), respectively [169-173] (Table 1.4). Subsequently, a number of additional activity-altering SNPs have been described, namely 190C>T (*2*19*), 364G>A (*2*12D*), 411A>T (*2*5I*), 434A>C (*2*17*), and 499G>A (*2*10*). However the effect of these particular SNPs on the overall *NAT2* acetylation status in different populations has yet to be investigated comprehensively.

A total of 27 different SNPs, defining 52 *NAT2* alleles, have thus far been described (http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/). Combinations of 1-4 SNPs, are classified into several haplotype groupings (allele clusters), which by popular convention, are usually classified via one of the reduced-activity SNPs defined above. However, as indicated in table 1.4, some of these SNPs also occur in differently defined haplotypes, for example the 341T>C SNP, whilst others, such as the 364G>A and 411A>T SNPs, occur solely in particular haplotypes. Hence these SNPs may be confounding factors during *NAT2* acetylation analyses, both when assigning the particular *NAT2* genotype, as well as when defining the overall acetylation phenotype.

Table 1.4. *NAT2* haplotypes and their defining-SNPs conferring reduced enzyme activity.

NAT2 Haplotype	Defining SNP	Amino Acid change[§]	Alleles containing the haplotype-defining SNP
2*5	341T>C	I114T	*2*5A – 2*5M*; *2*14C, 2*14F*
2*6	590G>A	R197Q	*2*6A – 2*6L*; *2*5J, 2*14D*
2*7	857G>A	G286E	*2*7A - 2*7B*; *2*6I, 2*6J*
2*10	499G>A	E167K	*2*10*
2*14	191G>A	R64Q	*2*14A – 2*14H*;
2*17	434A>C	Q145P	*2*17*
2*19	190C>T	R64W	*2*19*
2*5I	411C>T	L137F	*2*5I*
2*12D	364G>A	D122N	*2*12D*

[§] = See List of Abbreviation for amino acid abbreviations.
Table constructed from the information found in the official NAT website.

The 27 known *NAT2* SNPs are listed in Table 1.5. The majority of these SNPs (22/27) code for non-synonymous amino acid changes, whilst the remainder, namely, 111T>C, 282C>T, 345C>T, 481C>T, and 759C>T, encode synonymous amino acids, which by convention, are considered to impart an enzyme activity equivalent to the *NAT2*4* allele. However, Cascorbi *et al* [174], using caffeine as the probe substrate, found that the 282C>T SNP caused a reduced acetylation activity in a cohort of German individuals, thereby highlighting the significance of SNP validation, both genetically and phenotypically. Interestingly, the majority of the SNPs depicted in Table 1.5 are present in alleles that form sub-categories of the central *NAT2* haplotypes (*2*5*, *2*6*, *2*7*, and *2*14*), yet the effect on the NAT2 acetylation status for many of these SNPs is still undefined.

Table 1.5. *NAT2* SNPs and their effect on the Acetylation Phenotype.

Nucleotide Change	Amino Acid Change	Acetylation Phenotype	Defining Alleles
None	**Reference Sequence**	**Rapid**	***2*4***
29C>T	I10T	Unknown	*2*12J*
70T>A	L24I	Unknown	*2*5L*
111T>C	F37F[§]	Unknown	*2*6D*
152G>T	G51V	Unknown	*2*6M*
190C>T	R64W	Slow	*2*19*
191G>A	R64Q	Slow	*2*14*
203G>A	C68Y	Unknown	*2*5O*
228C>T	Y76Y	Unknown	*2*12I*
282C>T	Y94Y[§]	Rapid	*2*13*
341T>C	I114T	Slow	*2*5*
345C>T	D115D[§]	Unknown	*2*6L*
364G>A	D122N	Slow	*2*12D*
403C>G	L135V	Unknown	*2*12H*
411A>T	L137F	Slow	*2*5I*
434A>C	Q145P	Slow	*2*17*
458C>T	T153I	Unknown	*2*21*
472A>C	I158L	Unknown	*2*5N*
481C>T	L161L[§]	Rapid	*2*11*
499G>A	E167K	Slow, substrate dependant	*2*10*
518A>G	K173R	Unknown	*2*6G*
578C>T	T193M	Unknown	*2*5P; 2*12E; 2*13B*
590G>A	R197Q	Slow	*2*6*
600A>G	E200E[§]	Unknown	*2*20*
609G>T	E203D	Unknown	*2*12G*
622T>C	Y208H	Unknown	*2*12F*
638C>T	P213L	Unknown	*2*6K*
683C>T	P228L	Unknown	*2*14H*
759C>T	V53V[§]	Unknown	*2*5F*
766A>G	K256E	Unknown	*2*6H*

Nucleotide Change	Amino Acid Change	Acetylation Phenotype	Defining Alleles
803A>G	K268R	Rapid	*2*12*
838G>A	V280M	Unknown	*2*5M*; *2*6I*
845A>C	K282T	Rapid	*2*18*
859Del	S287 Frameshift	Unknown	*2*5H; 2*11B*
857G>A	G286E	Slow	*2*7*

Table 1.5 continued…. adapted from http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature

§ - indicates Synonymous amino acid changes, most of which have yet to be characterised phenotypically.

SNPs highlighted in colour represent the activity altering SNPs most often analysed in *NAT2* genotyping studies.

1.3.6.2.1 Functional impact of *NAT2* Polymorphisms

The 191G>A, 341T>C, 590G>A, and 857G>A SNPs, reduce the acetylation activity of the NAT2 enzyme, resulting in the slow acetylation (SA) phenotype [168, 175]. *In vitro* expression studies indicate that none of these SNPs alter the level of mRNA transcribed from these alleles [168, 176, 177], thereby indicating that they do not affect the rate of gene expression. However, the level of protein derived from the 191A, 590A, and 857A alleles was significantly reduced, whereas in the case of the 341A allele it was not [176]. Furthermore, only the 857A polymorphism exhibited an increased affinity towards the probe substrate (SMZ), as well as a decreased affinity for the cofactor acetyCoA. This latter finding supports the results of Hickman *et al* [178], who also reported that the 857G>A SNP exhibited an increased affinity for some substrates (SMZ and dapsone) but not others (INH and 2AF).

Similar *in vitro* functional characterisation of the other SNPs conferring the SA phenotype, namely, 190C>T (*2*19*), 364G>A (*2*12D*), 411A>T (*2*5I*), 434A>C (*2*17*), and 499G>A (*2*10*), also confirmed that their reduced acetylation capacity was due to the decrease in the level of the active protein, rather than to a change in the level of mRNA expressed from these *NAT2* alleles [175, 177, 179, 180]. In the case of the *2*12D* (364G>A and 803A>G SNPs) and *2*5I* (341T>C, 411A>T, 481C>T and 803A>G SNPs) alleles, the presence of the 364A and 411T SNPs respectively, affected the overall acetylation, resulting in undetectable levels for the corresponding proteins, when assayed via Western Blot analysis [180]. Thus the rapid acetylation phenotype, conferred by the 481C>T and 803A>G SNPs [168, 181] is completely abrogated in these alleles.

These authors also found that the 499G>A and 857G>A SNPs caused a moderate reduction in enzyme activity, approximately 40% lower than that observed for the reference *2*4* allele. In addition, the 499G>A SNP appears to affects the stability of the protein without compromising the kinetic activity of the *2*10* allele, with respect to the substrate and the cofactor acetyl-CoA. It would therefore appear that separate mechanisms may prevail in influencing the slow acetylator activity at the NAT2 protein level. This would imply that there could be different rates of slow acetylation activity, intrinsic to a particular SNP or haplotype, expressed via aberrant or improper protein folding, or inherently unstable proteins, that undergo an accelerated degradation within

the cellular environment. This kind of differential phenotypic activity has been demonstrated for the *2*6A*, *2*7B* and *2*14A* alleles, which exhibited a lower activity than found for the *2*5A*, *2*5B* and *2*5C* alleles [174, 182]. The activity of these *2*5* alleles was reduced by approximately 25% when compared to the *2*4* allele, whereas the activities for the *2*6A*, *2*7B* and *2*14A* alleles were reduced by more than 70% with respect to 2*4 over the same time period.

The determination of the individual's *NAT2* acetylation status generally focuses on an analysis of the following 7 SNPs as the means of defining the acetylation status, namely 191G>A, 282C>T, 341T>C, 481C>T, 590G>A, 803A>G, and 857G>A [168, 175, 177]. Whilst there is some contention surrounding the activity conferred by the 282C>T SNP, its inclusion is based upon the observation that it often occurs in linkage equilibrium, as part of the *2*5*, *2*6*, *2*7* and *2*14* allele clusters (haplotypes). Similarly, the 803A>G SNP is seldom observed as an individual SNP, and also forms part of the above haplotypes, where, by convention, they are considered to play a secondary role to the activity-altering SNP defining these clusters [177].

Furthermore, the 341T>C, 590G>A, and 857G>A SNPs also form part of the *2*14C*, *2*5J* and *2*6I* alleles respectively, where their role in the overall acetylation status is still obscure, since the phenotype of these alleles has not, to our knowledge, been determined via *in vitro* expression analyses. In addition, many reports in the literature genotype the 481C>T SNP as a proxy for the 341T>C SNP, since these SNPs are considered to be in linkage disequilibrium [62, 164, 183, 184]. However, this type of assay would yield erroneous results, since the 481C>T SNP is not present in several of the alleles classified among the *NAT2*5* haplotypes, such as for example, *2*5C*, *2*5D*, *2*5E*. These results therefore highlight the importance of investigating the contribution of the various *NAT2* SNPs both individually, as well as within the context of their respective alleles, in order to assay the accurate acetylation phenotype within individuals. This could be particularly important since the different SNPs may not be equally distributed within different individuals and/or populations.

1.3.6.2.2 Ethnic Distribution of *NAT2* Polymorphisms

Analysis of the acetylation status in different populations indicate that the frequency of occurrence of fast and slow acetylators differs between population groups [185-187]. The phenotypic fast acetylation is conferred by the homozygous wild type F allele (FF; usually *2*4*, but increasingly *2*12* and *2*13* are also included as the genotypic equivalent), and slow acetylation by the homozygous S allele (SS; combinations of either of the *2*5*, *2*6*, *2*7*, or *2*14* genotypic alleles).

Genetic analyses on the other hand, also suggests a heterozygous FS genotype, equated to an intermediate acetylation status in studies of certain population groups [38, 63, 173]. However, in studies of Caucasian populations the FS genotype is often considered to confer the FA phenotype [62, 171]. This bi- and trimodal acetylation classifications have persisted in the literature, despite the statement by Weber that "acetylation is determined by two major alleles without dominance." [46]. Secondly, the *NAT2* gene has been mapped to chromosome 8p22 [84, 86], and no epigenetic effects, such as an altered methylation status, have been reported for *NAT2* alleles. Hence it is more likely that these alleles are equally expressed in the cell. Thus, there is very little support for the idea that the F allele is the dominant allele, thereby arguing against the assignment of the FS genotype to a FA rather than an IA status.

Table 1.6 depicts the *NAT2* allele frequency reported for different population groups. Currently the number and type of alleles investigated in most populations is incomplete, and several of the interesting *NAT2* alleles are often not determined. For example the 191G>A SNP (*2*14* allele) is often omitted from the analysis since it is considered to be an "African" allele [164]. However it has also been found in Brazilians [188], Caucasians [174], Emiratis [189], and Hispanics [185], as well as ethnic individuals from South India [172], thereby suggesting that it would be erroneous to omit it from the population analysis based upon the physical appearance of the subjects in the cohort.

The general trend seen in Table 1.6 indicates that the FA genotype frequency is 20-25% in Caucasians, compared to 60-70% in Oriental populations [190]. In other parts of the world, the picture is difficult to interpret since the extant information often consists of phenotype-derived data only. Patin *et al* [167, 190], in reporting on the evolutionary history of the *NAT* genes, indicate an extended spectrum of NAT2 acetylation phenotypes

across the world, and importantly, within the African continent. They compared the results of studies in Eskimos, Koreans, Japanese, Papua New Guineans, and Pygmies, which indicate that the frequency of slow acetylators in these populations is very similar, at approximately 20%. In Germans and Russians on the other hand, the frequency of slow acetylators is 50-60%, similar to that reported for Zimbabweans [191], and some South African population groups [155].

Studies of African populations also indicate a wide spectrum of slow acetylation, ranging from 40% in Yorubas in Nigeria [192], 10-18% in Pygmies, and 70% in Somalis [193], to greater than 80% in northern Sudanese [194]. This data therefore indicates that an acetylation study of "African individuals" should not be extrapolated to include Africans in general, since there is a clear shift in the frequency of SA within the continent, from west to east, and even from north to south. This observation may be of particular importance with regard to the global TB pandemic, given that Sub-Saharan African is one of the regions of the world that experiences the highest burden of TB disease (WHO Report, 2007). Given this inter-relatedness between the individual's inherent capacity for metabolising INH and the very high burden of TB disease within the region, there may be considerable merit to the suggestion that INH treatment should be individualised to achieve the greatest treatment efficacy.

Table 1.6 ***NAT2*** **Allele (Haplotype) frequencies in different population groups**

Population	Allele Frequency (%)					Citation
	NAT2*4	***NAT2*5***	***NAT2*6***	***NAT2*7***	***NAT2*14***	
European						
British	26	45	27	1.6	n.d.	[62]
Scottish	20	49	27	3.6	n.d.	[195]
German	23	47	28	1.3	n.d.	[174]
French	18	52	25	1	n.d.	[192]
Spanish	25	39	27	1.4	n.d.	[196]
Portuguese	21	43	33	2.7	n.d.	[197]
Turkish	46	27	23	4	n.f.	[184]
North American						
Caucasian (USA)	25	45	28	2	n.f.	[164]
African-American	36	30	22	2	9	[164]
Asian						
Chinese	62	5	19	14	n.d.	[198]
Korean	60	2	23	15	n.f.	[199]
Japanese	61	1.8	25	10	n.f.	[161]
Indian	23	37	32	1.5	n.f.	[200]
Thai	38	4	33	20	n.d.	[201]
South American						
Embera[a]	63	10	4	23	n.d.	[202]
Ngwabe[a]	74	2.5	0	23	n.d.	[202]
Amerindians[b]	51	25	6	20	n.d.	[203]
Brazilian	20	38	26	4	3	[188]
African						
Egyptian	21	50	26	3	n.d.	[171]
Iranian	43	32	19	6	n.f.	[204]
Gabonese	27	41	22	2	8	[192]
Dogons	13	30	47	3	7	[192]
Tanzanian	30	34	21	3	12	[191]
Zimbabwean	28	31	21	6	14	[191]
Venda	23	39	22	5	11	[191]
Xhosa	60	20	15	2	3	[205]
Coloured	54	25	11	3	5	[205]

n.d. = not determined by experimental assay; n.f. = not found in cohort studied
[a] = Native American tribes resident within Panama and Colombia
[b] = Native American tribes resident within Argentina and Paraguay

1.4 NAT Acetylation and TB treatment Efficacy

It is feasible that INH is exclusively inactivated by the NAT2 acetylation reaction, since NAT1 has not yet been shown to be involved in the acetylation of INH [97]. However, given the significant structural similarity between these enzymes [83], it is not inconceivable that NAT1 and NAT2 are isoenzymes, with the capacity to act synergistically. In this context, it is interesting to note that in 1951 Mark *et al* [206] reported that PABA (a NAT1-specific substrate) is formed during the metabolism of PRO (a NAT2-specific substrate), thereby respectively implicating both NAT2 and NAT1 in this metabolic pathway. However, the recent *in vitro* analyses by Kawamura *et al* [97] indicated that NAT1 and NAT2 exhibit a dissimilar substrate-specificity towards PABA. On the other hand, Fretland [207] *et al* showed that 4-aminobiphenyl (ABP) is acetylated by both Nat1 and Nat2, the murine homologues of human NAT2 and NAT1, respectively.

A study by Mitchell and Warshawsky [125] showed that the upstream promoter domains of both these genes contains promoter sites which are induced in a tissue-specific rather than substrate-specific manner. *in vitro* tests in a liver cell line, indicated that both PAS (NAT1-specific substrate) and SMZ (NAT2-specific substrate) induce similar rates of expression for both NAT1 and NAT2 enzymes, whereas in bladder cells, both these substrates only affected NAT1 expression. This is an interesting result, since it is commonly accepted that SMZ is a NAT2-specific substrate, which is not metabolised by NAT1 [97], and therefore the question remains as to whether these genes, given their close proximity, and similarity as isoenzymes, act exclusively and/or co-operatively in the acetylation metabolism. In a study conducted in Caucasians, Smelt *et al* [208] reported that the prevalence of the *NAT1*10* and *NAT2*4* haplotype occur 3.5 times more frequently than expected by chance association. Subsequent studies [147, 209] have reported a similarly higher occurrence for the *NAT1*10* allele in Caucasians genotyped as *NAT2* FA. The significance of any linkage disequilibrium between NAT1 and NAT2 is unknown, and has yet to be tested with respect to these and other NAT haplotypes in various diseases, if we are to develop a better understanding of the involvement of NATs in the overall acetylation reaction.

1.4.1 Optimising the INH Dosage

Over the past 50-60 years the treatment of TB has been refined to the current internationally standardised DOTs (directly observed treatment, short course) protocol, in which patients are treated for a period of six months, with a defined drug regimen containing INH. An essential component of the DOTs protocol is a well-structured programme to manage ambulatory patients, many of whom are socially disadvantaged. This requires that the strategy for optimal large-scale therapy be adapted to a single most suitable regimen for this type of programmatic care, since there is a logistical and economic advantage to this standardised care. At present the uniform INH dosage recommended by the World Health Organisation (WHO) is 5 mg/kg body weight, with a range of 4–6 mg/kg [210]. However, an argument can be made for individualised treatment, based upon the uptake and metabolic characteristics of individual patients, given the apparent localisation of the burden of disease in some parts of the world.

In developing countries such as South Africa, individualized treatment may not be practicable owing to very limited resources, and large numbers of patients. In this situation an equitable balance needs to be achieved between treatment efficacies on the one hand, and drug toxicity on the other. Adverse drug reactions (ADRs) are a well known feature of the TB treatment regimens, and it has been established that INH causes toxic side effects in a number of TB patients [211, 212]. In this respect the two primary forms of toxicity are peripheral neuropathy and hepatotoxicity. Hepatotoxicity, although not common amongst TB patients, can be fatal, and has been suggested to occur more frequently in slow acetylator (SA) individuals [213].

Treatment efficacy therefore, is dependant upon the combined effects of the anti-bactericidal activity occurring at a particular INH dosage, together with the individual's metabolism of INH. INH pharmacokinetic studies indicate that the maximum early bactericidal activity (EBA) of the drug occurs at 2 hours post dose, at a concentration of 2-3 µg/ml [173]. However, because the individual's inherent acetylator status would ultimately influence this response, it has been suggested that FA and IA adult individuals, due to their relatively rapid rate of acetylation, would exhibit a sub-optimal INH anti-bactericidal activity, when given the WHO-recommended doses of 250-300 mg of total INH. Therefore these individuals should arguably be given a larger dosage of INH to

achieve an optimal EBA, as well as reduce the risk of developing drug resistant *Mtb* strains [214, 215] resulting from doses of INH which may be sub-lethal for these mycobacteria. Whilst the clinical importance of these observations with regard to individuals who are fully compliant with the treatment regimen has been questioned [216], the deficiencies of FA and IA individuals may become more pertinent, since the lower INH concentrations may translate into treatment failure or development of mono-resistance, as well as progression to multi-drug resistance, and hence persistence of disease [217, 218].

When compared to RA individuals, SA take approximately twice as long to metabolise an equivalent amount of INH (Parkin, DP (1996) Ph.D. Dissertation. http://hdl.handle.net/10019.1/1267), and should probably be given a smaller INH dosage to reduce the risk of the development of drug-induced hepatotoxicity [211, 219], since it has been suggested that the SA may be more susceptible to this adverse effect [213]. Drug-induced hepatotoxicity (DIH) leads to serious deleterious consequences for the efficacy of TB treatment, since it necessitates interrupting the treatment regimen, and consequently the treatment period is prolonged. In additional, the interruption of treatment may also increase the risk for the development of drug-resistant organisms, which are more difficult to treat, and require the use of less effective, yet more toxic "second-line" anti-tuberculosis drugs.

The relationship between the acetylator status and the development of DIH is still controversial, mainly because the incidence of DIH is often idiosyncratic and difficult to predict [220]. However, some reports in the literature postulated that RA individuals would be more prone to DIH due to the rapid accumulation of toxic intermediates formed during the metabolism of drugs such as INH [221]. More recently SA have been implicated as being more prone to the development of ADR [213, 222]. This apparent contradiction suggests that the role of the individual's acetylation status requires further investigation to fully elucidate the contribution of RA and SA to DIH.

1.4.2 The role of TB drugs in Hepatotoxicity

The WHO recommended treatment for drug susceptible TB consists of a drug regimen of isoniazid, rifampicin (RIF), pyrazinamide (PZA), and ethambutol (EMB) maintained for a period of at least 6 months (WHO, Treatment of Tuberculosis; guidelines for national programmes, 2009). Despite the effectiveness of this regimen in curing more than 95% of patients of their tuberculosis [223], it is well-known that DIH occurs within the first 2 - 4 months of treatment, which can severely reduce efficacy of treatment [224]. Anti-TB drugs are usually co-administered in combination therapy, and whilst the contribution of individual drugs to the development of DIH is still controversial [225], it has been shown that INH, RIF and PZA are potentially hepatotoxic [226, 227]. In this regard most of the data involves the well known metabolism of INH, since this drug is widely used as a single drug regimen for the prophylactic treatment of latent TB [228]. Currently it is believed that the drug-induced hepatotoxicity is caused by toxic metabolites formed during the metabolism of INH (Fig. 1.10).

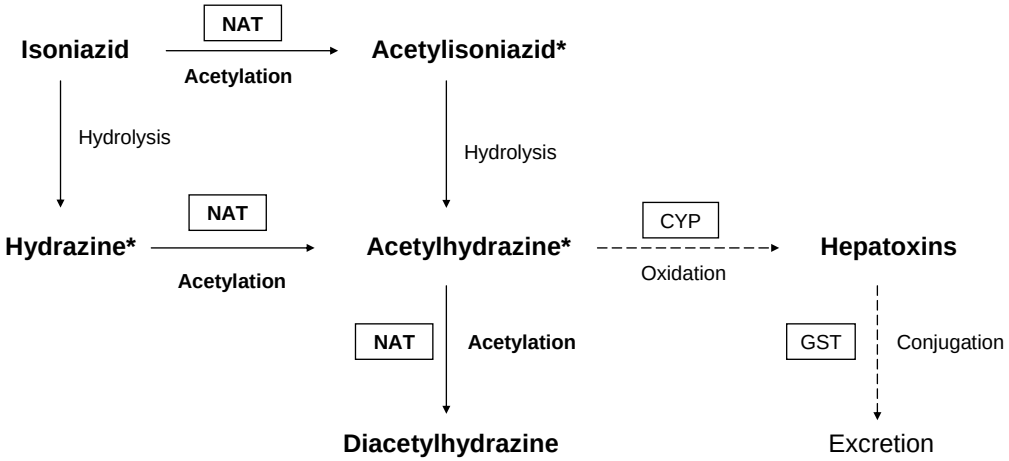


Figure 1.10. Proposed route of isoniazid metabolism in the human liver. Drug metabolising enzymes involved in the pathway are indicated in text boxes, whilst the major catalytic pathways are highlighted by solid lines. Potential, subsequent reactions are indicated by dotted lines. Hepatotoxic compounds are indicated by the asterisk.
Adapted from Yew *et al*, (2006) [229] and Tostman *et al*, (2008) [225].
NAT = arylamine *N*-acetyltransferases; CYP = Cytochrome P450 enzymes; GST = Glutathione *S*-transferase

The primary hepatotoxic intermediates formed during INH metabolism have been identified as acetylhydrazine and hydrazine [226] formed via the metabolic pathways outlined in Fig.1.10. Initially INH is acetylated by NAT2 to form acetylisoniazid, which is subsequently hydrolysed to acetylhydrazine [230]. The acetylhydrazine then undergoes another round of acetylation, forming the non-toxic metabolite diacetylhydrazine. In addition, the acetylhydrazine can be oxidised into hydrazine via an alternate pathway mediated by the cytochrome P450 (CYP450) family of enzymes, which appear to be inducible by RIF [229]. Furthermore, whilst it is believed that the NAT2-mediated acetylation of INH is the predominant metabolic pathway, the hydrolysis reaction to form hydrazine [231], indicates that RIF and INH can have a combined effect on hepatotoxicity [232]. This latter reaction has recently been shown to proceed via the RIF-mediated induction of CYP2E1 [233].

Thus far, the incidence of DIH has been investigated mainly in Caucasians and certain Asian populations, and shown to vary between 2% and 28% [225, 234]. The direct contribution of RIF and/or PZA to this hepatotoxicity is still largely unknown, yet PZA toxicity has been shown to be largely dose-related, and a low dosage of 20-30 mg/kg of PZA obviates the development of acute DIH [226]. The metabolism of PZA proceeds via xanthine oxidase, and currently the mechanism of PZA-toxicity is controversial: PZA was shown to inhibit CYP450 enzymes in a study of rat microsomes [235], whereas PZA had no effect in a similar study using human liver microsomes [236]. However, it has recently been shown that PZA toxicity can be induced by hydrazine *in vitro* [237]. This latter result, together with the proposed RIF-mediated formation of hydrazine, could suggest a potential mechanism whereby INH-induced hepatotoxicity can be exacerbated by both RIF and PZA during treatment.

Mycobacterium tuberculosis is a slow growing organism, and effective anti-TB treatment regimens are of necessity relatively long, and often protracted. Consequently, treatment efficacy is constrained by the development of DIH, leading to treatment suspension or interruption, which further concatenates the treatment period, and could facilitate the emergence of drug-resistant organisms, thereby fuelling the TB pandemic.

1.4.3 The Effect of NAT Acetylation on Treatment Efficacy

1.4.3.1 Acetylation Polymorphisms and Treatment Response

TB treatment efficacy hinges upon the effective elimination of *Mtb* organisms from the human host [14], yet the activity of INH, the primary anti-TB drug, is diminished by the human NAT2 enzyme [18, 35, 39]. The enzyme's activity, in turn, is defined by several well-characterised exonic SNPs [44, 50, 62], which collectively constitute the mechanism underlying the "classical" acetylation polymorphism, of fast (FA), intermediate (IA) and slow acetylation (SA). Hence, the overall acetylation activity will be differently constrained by the individual alleles defining the slow (SS), intermediate (FS), or fast (FF) genotype (or diplotype). This activity can primarily be correlated with the SNPs at 191, 341, 590 and 857, which characterise the *2*14*, *2*5*, *2*6* and *2*7* alleles respectively [140, 168]. Secondly, it has been demonstrated that these individual SNPs, do not affect the rate of acetylation to the same degree [176, 182]. So for example, Grant *et al* [182] demonstrated that the *2*5* allele still exhibited 75% acetylation activity when compared to the reference *2*4* allele, whereas the activities for the *2*6A*, *2*7B* and *2*14A* alleles dropped to less than 22% of the *2*4's* activity over the same time period. Similarly, Zang *et al* [177] found that the enzyme encoded by the 857G>A SNP exhibited a greater binding affinity for acetyl-CoA, which indicates that the formation of the activated acetyl-NAT intermediate (Fig.1.2) is enhanced in the *2*7* allele. Furthermore, this particular NAT2 variant exhibited distinct binding affinities for different substrates.

These results therefore suggest that the metabolism of INH would be subjected to varying degrees of acetylation activity, depending upon the identity of the individual SNPs and alleles in the genotype. Hence combinations of the "reference" allele with different polymorphic alleles, will give rise to differential rates of acetylator activity amongst IA individuals, and similarly, different polymorphic allele combinations would give rise to differential rates of SA activity. In the context of treatment efficacy, these differential rates of acetylation between FA, IA and SA individuals, could translate into individualised serum INH concentrations in patients resulting in degrees of treatment response and efficacy, which may form the basis for the varying treatment responses observed in patients when employing the recommended doses of INH [217, 218].

As indicated previously, there is a direct concordance between acetylator genotype and phenotype, and the frequency of the activity-altering SNPs differs amongst population groups. Thus for example, the *2*5* allele predominates amongst Caucasians, the *2*6* allele is more prevalent in Asian populations and amongst Amerindians in South America, the *2*7* allele predominates (see Table 1.6). Furthermore, Grant *et al* [182] has shown that the overall activity of the NAT2 proteins derived from individual SNPs, is also very dissimilar. For example, a comparison of the acetylation half life between the enzymes encoded by the *NAT2*5* and *NAT2*6* alleles, showed that the overall rate of acetylation was reduced to a greater extent in the case of *NAT2*6*, when compared to the wild-type *NAT2*4* allele. Thus, depending upon the relative frequency of particular *NAT2* SNPs, the relative rate of acetylation for substrates like INH would vary between populations. In this context, the overall *NAT2* genotype could determine the different responses to treatment observed between populations, as well as between individuals with different genetic backgrounds within a population.

The information available on *NAT2* genotypes for populations on the African continent is generally very limited, and the SNP frequencies, contributing to the acetylation polymorphism have not been well documented [155, 190, 191, 205]. Similarly, whilst the aetiology and incidence of DIH is relatively unknown, a lack of treatment efficacy due to DIH has variously been linked to old age [238], malnutrition [239], gender [240] and alcohol intake [241]. However, these research publications have also implicated the role of inherent genetic factors, and recently these aspects have received more attention.

Several reports describe an association between anti-TB drugs and hepatotoxicity, particularly in relation to INH and its known metabolism [221, 224-226, 228]. This serves to underline the importance of delineating the effects of the acetylation polymorphism in countries with a high burden of TB disease. Thus DIH and non-compliance with treatment may be fuelled by the activities of the inherent drug metabolising enzymes, and hence act synergistically in reducing treatment efficacy, and facilitate the continued spread of disease. The overall effect would be an increase in the morbidity and mortality among patients, whilst also facilitating the emergence of drug resistant organisms.

The polymorphic acetylation status inherent in individuals requires further investigation to determine its contribution to, and association with, the epidemiology of the global TB

pandemic, given the fact that TB is an infectious disease. In this respect, the metabolism of therapeutic drugs like INH may affect various aspects of treatment efficacy either separately and/or additively. Therefore, the general effect of acetylation in SA and RA TB patients in diverse populations, from high burden countries, may need to be investigated more comprehensively, to determine the contribution of population-specific activity-altering SNPs, in reducing the optimal serum concentration of INH during treatment. Secondly, the role of the formation of toxic metabolites such as hydrazine (Fig.1.10), suggested to play a greater quantitatively role in the DIH in SA rather than RA [242] is still largely unknown, and needs further investigation, particularly in the developing countries in Sub-Saharan Africa, where an analysis of the frequency of *NAT2* SNPs, as they pertain to the acetylation polymorphism has been neglected. In this context, it is interesting to note that the prevalence of the acetylation polymorphism conferred by the NAT2**6* allele in Japanese cohorts, has been associated with an increased risk of DIH in SA rather than RA individuals. [222, 243]

During the lucrative spice-trade era between West and East, South Africa was an important stop-over destination, and consequently its populations have assimilated genetic ancestries from indigenous African, Caucasian and Asian populations. This is particularly evident in the local population group known as South African Coloured (SAC) [244, 245]. Given the current information with regard to the acetylation polymorphism in Caucasian and Asian populations, as well as the dearth of information in African populations, South African populations such as the SAC may afford a seminal opportunity to investigate acetylation polymorphisms against the inherent genetics in local population enclaves. The primary significance of such an analysis would afford a better understanding of the assimilated acetylation knowledge, as well as contribute to our understanding of the epidemiology of TB, particularly in a high burden of disease-setting, such as the Western Cape region of South Africa.

1.5 Hypothesis and aims of the study

The Arylamine *N*-acetyltransferase enzymes, NAT1 and NAT2 are important drug metabolising enzymes because they acetylate various xenobiotics and key therapeutic drugs, such as p-amino-salicylate [44] and isoniazid [18], used in the treatment of tuberculosis [35, 39, 75].

These enzymes exhibit an inter-individual variation in this acetylation metabolism due to several well characterised genetic polymorphisms. The efficacy of anti-tuberculosis treatment regimens may therefore be compromised by the activity of these NAT enzymes. Furthermore the inter-individual variation in enzyme activity may also be effected by differences in the frequency distribution of these genetic polymorphisms within admixed populations.

1.5.1 Aims of the study

- To analyse the population diversity of *NAT* polymorphisms in adults in local ethnic groups.

- To investigate *NAT* polymorphisms in a South African Coloured community experiencing a high incidence of tuberculosis disease.

- To test the effect of *NAT2* polymorphic variation in children diagnosed with tuberculosis disease.

- To evaluate *NAT2* acetylation polymorphisms in children to establish effective treatment guidelines for isoniazid dosing.

- To evaluate the influence of the inter-individual variation in the *NAT2* acetylation capacity on the early bactericidal activity of isoniazid.

CHAPTER 2

Materials and Methods

2.1 Chemicals and Reagents

All chemicals utilised in the experimental procedures were of analytical grade quality and purchased from the British Drug House (BDH) Laboratory Suppliers (UK), Merck (Germany), and the Sigma-Aldrich Company (USA).

Molecular biology reagents, restriction enzymes, and kits for the manipulation of DNA were purchased from Invitrogen (USA), the Promega Corporation (USA), Roche Molecular Diagnostics (Germany), Fermentas International Incorporated (Canada) and Amersham Biosciences (General Electric Healthcare, USA).

The solutions, buffers, reagents and laboratory equipment, as well as the resources for the manipulation of the data, are detailed in the text when required and also at the end of this chapter.

2.2 Sample Cohorts

Unaffected and unrelated individuals from among the Caucasian, Coloured and Xhosa ethnic population groups, resident in the Western Cape Province of South Africa, were selected for inclusion in the study. These individuals were classified on the basis of their own self-identification with these ethnic groups, based upon the collective values and cultural norms ascribed to by the individual group. Individuals from the Xhosa group, although an established ethnicity in the Western Cape province, often still identify with a group cultural home-base in the Eastern Cape region of South Africa. The individuals drawn from the Caucasian and Coloured groups consider themselves to be an endemic population to this region of the country.

Coloured and Xhosa TB patients' samples were sourced from the Tygerberg and Brooklyn Chest Hospitals, as well as from suburban clinics within the Ravensmead-Uitsig area of the Cape Town metropolitan region.

2.3 Sample Preparation and Analysis

2.3.1 DNA Purification

Genomic DNA (gDNA) was prepared from 3-5 millilitres (ml) of peripheral whole blood collected from each individual, in separate EDTA-containing vacutainer tubes (BD Vacutainer Systems, Plymouth, UK). The blood samples were collected on ice, and transported to the laboratory on the same day, where they were then stored at –20 degrees Celsius (°C). This approach enabled us to obtain good yields of deoxyribonucleic acid (DNA), from each blood sample; a minimal complemented of ribonucleic acid (RNA) was also obtained, during each purification.

Good quality high molecular weight gDNA was extracted using a salting out procedure [246]. The method entails the removal of the cellular proteinaceous material by simple dehydration and precipitation, which obviates the use of hazardous organic solvents.

Each blood aliquot was diluted with 5-6 volumes of ice-cold Cell Lysis Buffer (320 mM sucrose, 1% Triton X-100, 5 mM $MgCl_2$, 10 mM Tris, pH 7.6) in a sterile 50 ml polypropylene centrifuge tube (Corning Incorporated, USA), and then mixed vigorously by hand-inversion 20-30 times. Thereafter the tubes were centrifuged at 3000 revolutions per minute (rmp) for 15 minutes at 4°C, in an Eppendorf centrifuge, model 5810R. The resulting pellet was resuspended in 6 ml of Nuclei Lysis Buffer (400 mM NaCl, 10 mM Tris, pH 8.2, 2 mM EDTA, pH 8.2), supplemented with Sodium dodecyl sulphate (SDS) to a final concentration of 1%, and digested with 200 micrograms (µg) of proteinase K enzyme for 2 hours at 55°C (concentration of the proteinase K stock solution is 10 milligrams per millilitre (mg/ml)).

Thereafter, 2 ml of a saturated sodium chloride (NaCl) solution (> 6 M) was added to the sample tube, and the contents gently hand-mixed to homogeneity, to reduce shearing of the DNA. The tubes were then centrifuged at 4000 rpm for 30 minutes at 4°C, to pellet the proteinaceous debris. gDNA was precipitated from the supernatant by the addition of 2.5volumes of room temperature absolute ethanol. The precipitated gDNA was allowed to float to the meniscus, before being spooled out with a Pasteur pipette, the end of which had been heat sealed.

The purified gDNA was washed with 70% ethanol, briefly heat dried at 37°C (15-20 minutes), and then re-dissolved in 800 µl of sterile 0.5x TE (10 mM Tris, pH 8.0, 1 mM EDTA, pH 8.0) buffer [247] for 16 hours at 4°C.

2.3.2 DNA Concentration and Purity

The quantification of the purified nucleic acid solution was determined spectrophotometrically at 260 and 280 nanometres (nm), by measuring the absorbance (Abs) of a diluted sample aliquot placed in a beam of light having a 1centimetre (cm) path length. Under these conditions a solution of double stranded DNA with a concentration of 50 µg/ml has $A_{260} = 1$ [247].

The ratio between the readings at 260 nm and 280 nm (A_{260}/A_{280}) - used to estimate the purity of the gDNA in solution - was generally found to have a value between 1.8 and 2.0, thereby demonstrating good deproteinisation of the samples. Solutions of pure DNA have an A_{260}/A_{280} value of 1.8 [247]. In those samples where the value of the A_{260}/A_{280} ratio was closer to 2, the integrity of the gDNA was evaluated by gel electrophoresis against a commercially supplied preparation of purified Lambda DNA (Roche, USA). A calculated 500 ng quantity of gDNA was assessed against 500 ng of uncut lambda DNA. The samples were separated in 0.7% Seakem® LE agarose (BioWhittaker Molecular Applications, USA (BMA)), stained with a fluorescent DNA dye (GelStar® nucleic acid stain (BMA), and visualised under ultra-violet light. In all cases the gDNA appeared as a single high molecular weight band, comparable in size and intensity to the band of lambda DNA. Furthermore these samples generally exhibited a minimal complement of RNA, apparent as a lesser staining front lower down the gel.

2.3.3 Agarose Gel Electrophoresis

Agarose gels, utilising either Seakem® LE agarose (BMA) or MetaPhor® agarose (BMA) were used for the electrophoresis of gDNA, as well as to assess the products of PCR amplification. PCR product sizes were compared to commercially supplied DNA markers – 100 bp DNA Ladder (Promega, USA), or GeneRuler 100 bp Plus (Fermentas

Inc., USA). The gels were prepared in 0.5x Tris-Borate-EDTA (TBE) buffer (45 mM Tris, 45 mM Borate, 1 mM EDTA, (pH 8.0)), in a mini-gel apparatus (Omeg Scientific, South Africa) with gel dimensions of 100 mm x 65 mm x 3.5 mm, or when required for the analysis of many samples, in the larger GNA 200 submarine unit (Pharmacia, USA), with gel dimensions 200 mm x 200 mm x 4 mm.

Sample DNA aliquots were mixed with a Blue/Orange DNA loading dye (Promega, USA), loaded onto the gel, and then electrophoresed at 15 volts per centimetre (v/cm), using 0.5x TBE as conductance medium. Thereafter the gels were stained at room temperature for 20 minutes, in a solution of 0.5x TBE containing the GelStar® Nucleic Acid stain (BMA), diluted to a 1x final concentration. The gels were visualised on the UV transilluminator (UVP, USA), and photographed using the Kodak Electrophoresis Documentation and Analysis 120 System (Kodak Digital Science, Eastman Kodak Co., USA).

2.3.4 PCR Amplification and Analysis of the *NAT2* Gene

A fragment of the *NAT2* gene (http://www.ncbi.nlm.nih.gov/nuccore/219871) was amplified via the Polymerase Chain Reaction (PCR) technique (Roche, USA) using the gene-specific primer set HuNAT14 (5'-^{726}GACATTGAAGCATATTTTGAAAG^{748}-3') and HuNAT16 (5-^{1724}GATGAAAGTATTTGATGTTTAGG^{1702}-3'), previously described by Hickman and Sim [62]. All primers were supplied by Integrated DNA Technologies (www.idtdna.com). The PCR amplification reaction was performed in a total volume of 100 µl, and consisted of 150 ng of gDNA, 20 mM Tris buffer (pH 8.4), 50 mM KCl, 0.2 mM each of dATP, dGTP, dCTP, and dTTP, 2.5 mM $MgCl_2$, 0.25 µM of the HuNAT14 and HuNAT16 primers, and 1 unit of *Taq* DNA Polymerase I (Gibco BRL Life Technologies, USA). The amplification reactions were carried performed in the Gene Amp PCR System 2400 (Applied Biosystems, USA) which was programmed to execute the following temperature cycling conditions:- initial denaturation at 94°C for 5 minutes, followed by 35 cycles of alternating denaturation at 94°C for 30 seconds, annealing at 56°C for 40 seconds, and polymerisation at 72°C for 80 seconds; this was followed by an extended polymerisation at 72°C for 10 minutes, where after the sample tubes were cooled to 4°C, and maintained at that temperature for 30 minutes.

A 5 µl aliquot of each sample reaction was analysed for *NAT2* product specificity and PCR efficiency by submarine gel electrophoresis (described above) using the Omeg mini gel apparatus. In addition, the specificity of the PCR product was assessed by the restriction enzyme analysis with *Hind*III and *Hinc*II (Fig. 2.1).

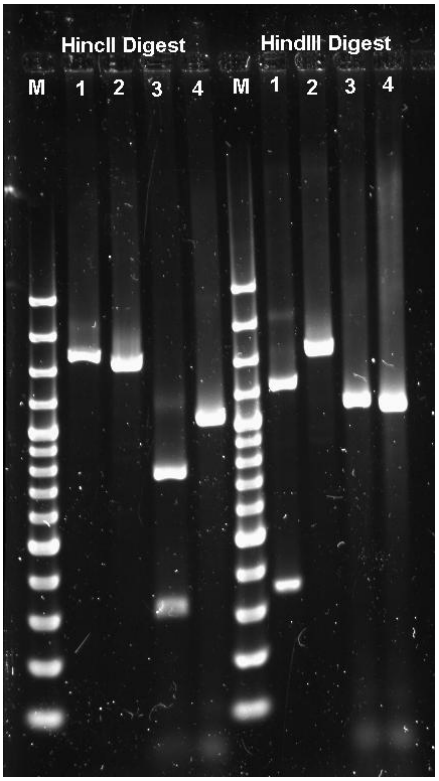


Figure 2.1 *NAT1* and *NAT2* amplified products restricted with *Hinc*II and *Hind*III enzymes, to indicate the specificity of the reactions. NAT1 (lane 1) is only cleaved by *Hind*III, whilst NAT2 (lane 3) is only cleaved by *Hinc*II. Lanes 2 (NAT1) and 4 (NAT2) represent undigested PCR product. Lane M contains the GeneRuler^{TM} 100bp DNA ladder, (SM0321, Fermentas International Inc., Canada), with 14 bands decreasing in size from 3000 bp, 2000 bp, 1500 bp, 1200 bp, 1000 bp, 900 bp, 800 bp, 700 bp … to 100 bp.

2.3.5 Restriction Endonuclease Analysis of the *NAT2* gene

A 15 µl aliquot of each 100 µl PCR sample solution were digested separately with the restriction endonucleases *Bam*HI, *Dde*I, *Fok*I, *Kpn*I, *Msp*I, and *Taq*I, respectively. All restriction endonuclease (R.E.) digestions were carried out in accordance with the recommendations of the specific R.E. manufacturer. *Bam*HI, *Kpn*I, *Msp*I, and *Taq*I restriction enzymes were supplied by Fermentas (Canada), *Dde*I was supplied by Roche (Germany), and *Fok*I was supplied by Amersham Biosciences (USA).

Restriction endonuclease incubations were carried out at 37°C, - except for *Taq*I, which was carried out at 65°C for 3-4 hours utilising 1unit of enzyme per 10 ul aliquot of the PCR reaction volume. In the case of the *Fok*I restriction, the digestion period was limited to a maximum of 2 hours, and the enzyme quantity was reduced to 0.6 units of enzyme per 10 µl PCR aliquot. This enzyme is prone to non-specific cleavage of the DNA in situations of over digestion - that is, longer than 2 hour incubation and/or a ratio of greater than 5 units *Fok*I per µg DNA [248].

An additional restriction of the *Bam*HI, *Kpn*I and *Msp*I restricted products was also performed using *Pst*I. This R.E. cleaves the *NAT2* PCR product twice at nucleotide positions 324 and 621 of the amplified sequence, generating cleavage products of 377 bp, 324 bp and 297 bp (Fig. 2.2). These latter cleavage products provide additional information to score the *Bam*HI, *Msp*I and *Kpn*I digestion products, respectively.

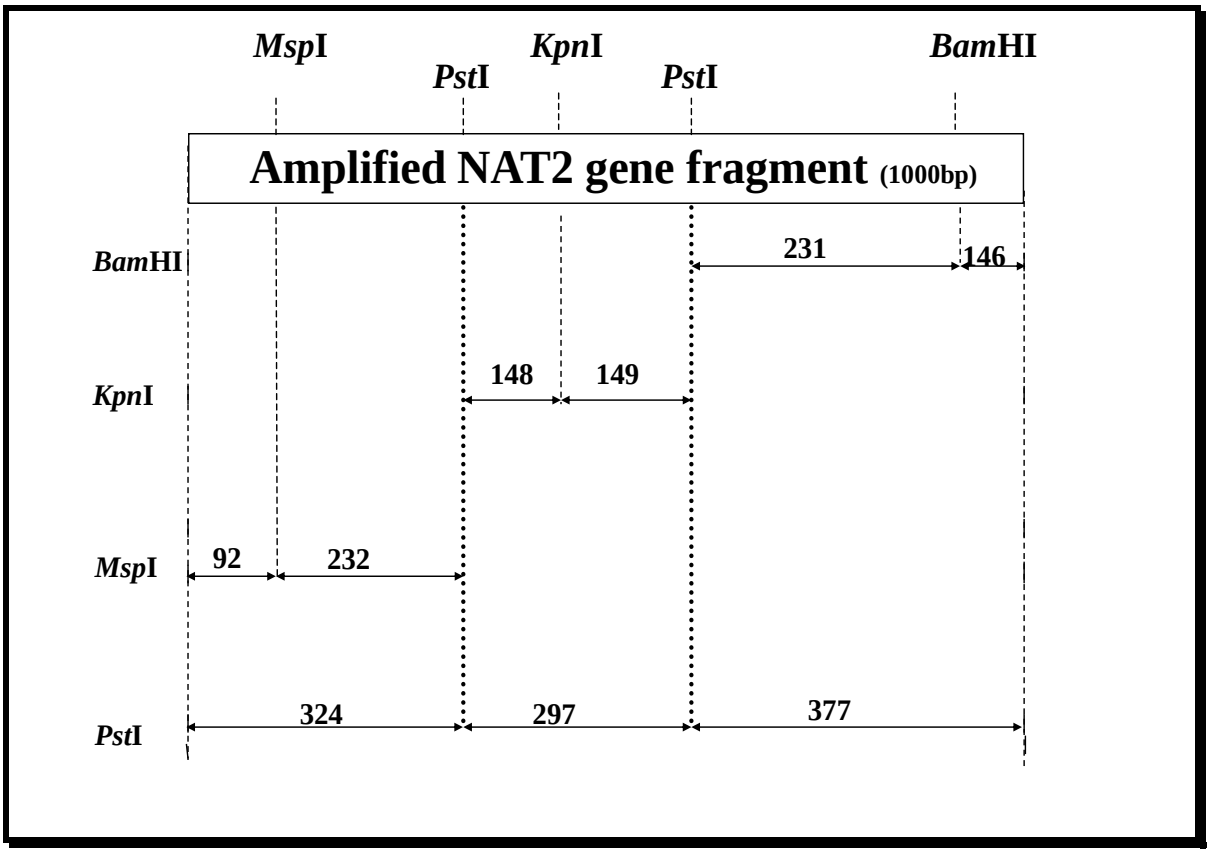


Figure 2.2 Diagram showing the relative restriction enzyme cleavage sites in the *NAT2* gene product. The additional cleavage with the *PstI* enzyme (round dotted lines) generates fragment lengths which facilitate the *Msp*I (324bp), *Kpn*I (297bp) and *Bam*HI (377)-specific analysis via polyacrylamide gel electrophoresis.

These double digests generate fragment profiles possessing bands of less than 500 bp in size, which are specific for *Msp*I, *Kpn*I and *Bam*HI, respectively. This facilitates the easy and accurate resolution of these profiles via non-denaturing polyacrylamide gel electrophoresis, which improves the sensitivity of the analysis (Fig. 2.3). The *Pst*I restriction sites, which occur at nucleotide positions 1050 and 1347 of the published *NAT2* gene sequence (http://www.ncbi.nlm.nih.gov/nuccore/219871) were found to be conserved in all of the samples analysed, and no polymorphic sequence variation have been observed in our sample cohorts using this enzyme.

In contrast to the above, the amplified *NAT2* gene sequence possesses several cutting sites for each of the *Dde*I, *Fok*I and *Taq*I R.E.s, generating enzyme-specific profiles that are also conducive to analysis via non denaturing polyacrylamide gel electrophoresis.

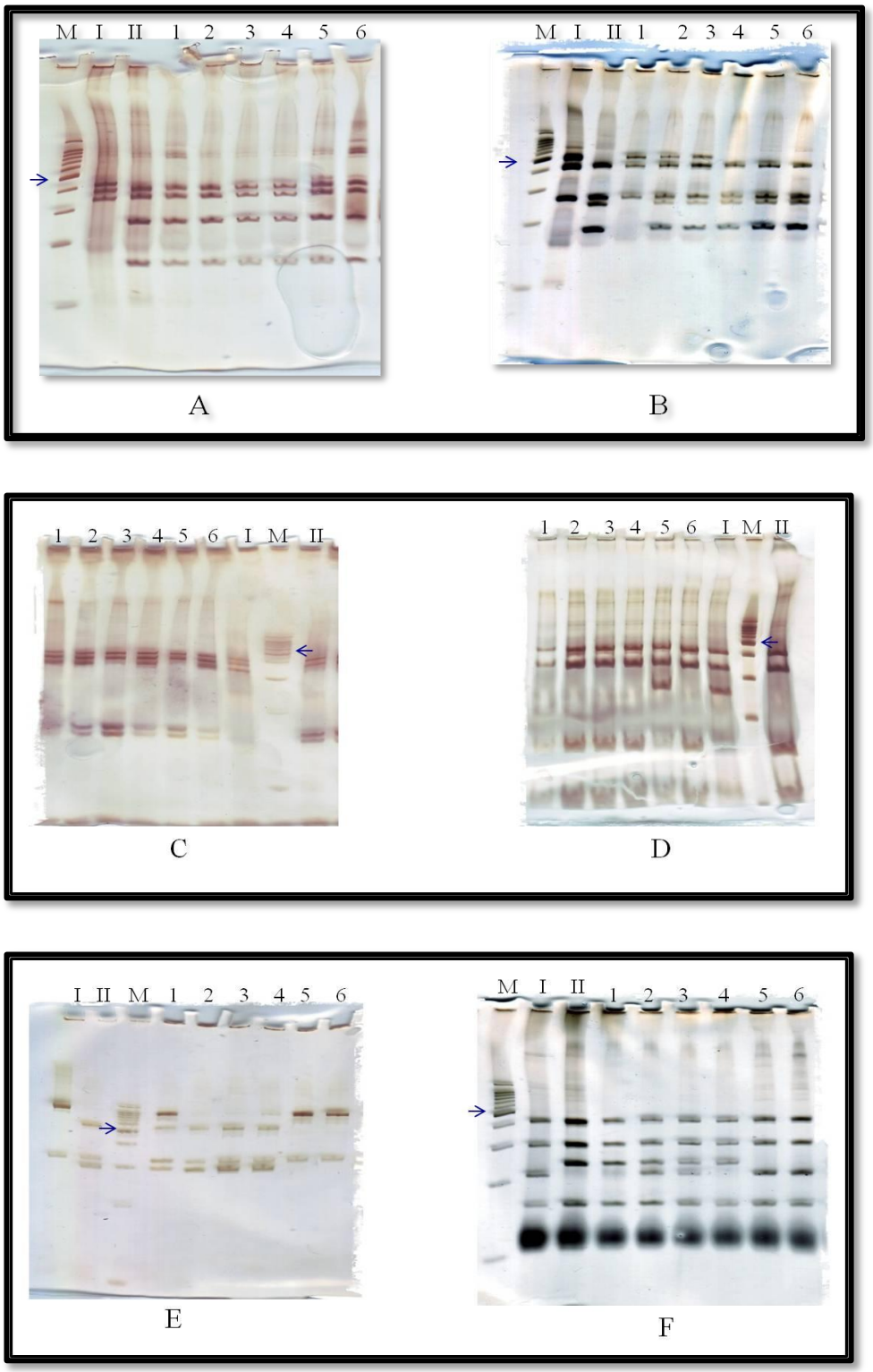


Figure 2.3 Restriction enzyme profiles for *Bam*HI-*Pst*I, *Taq*I, *Kpn*I-*Pst*I, *Msp*I-*Pst*I, *Fok*I, and *Dde*I cleavage, (panels A, B, C, D, E, and F, respectively). Lane M: 100bp DNA molecular weight marker ladder (Promega, USA), with the 500bp size marker indicated by the arrow. Lanes I and II are control sample digestions for each particular digestion series. Lanes 1 - 6 represent the results for individual test samples for each digestion series.

2.3.6 Polyacrylamide Gel Electrophoresis

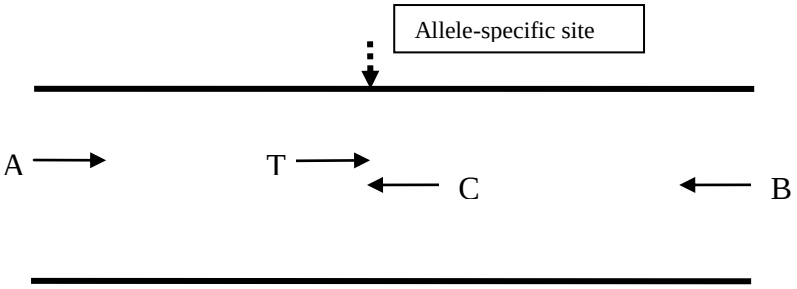
The *NAT2* DNA profiles generated with specific R.E.s were resolved in a Mini-PROTEAN® II gel electrophoresis system (BioRad Laboratories, USA) using a 5% non-denaturing polyacrylamide gel. The gels (102 mm x 73 mm x 0.75 mm) were prepared in accordance with the method detailed in Current Protocols in Molecular Biology [249]. A commercially supplied 40% stock solution mixture of acrylamide/bis-acrylamide (19:1) (Sigma- Aldrich, USA) was used in the preparation of these gels. The gels were allowed to polymerise at room temperature for 1-2 hours prior to being used. Gels were run at 100 volts for 1 hour using 1x TBE buffer as the electrophoretic buffer.

After electrophoretic separation, the DNA bands were visualised and fixed by silver staining [250]. The gel was washed in 50ml of 7.5% acetic acid solution for 10 minutes at 37°C, after which they were rinsed in water for 3 minutes at 37°C. This washing step was repeated thrice. Next the gel was stained for 20 minutes at 37°C with 50 ml of 0.1% silver nitrate solution, containing 0.056% formaldehyde, followed by a very brief wash in water for 10 seconds. The DNA band profiles were developed by incubating the gel in 50 ml of ice-cold 3% sodium carbonate solution, supplemented with 0.056% formaldehyde, and sodium thiosulphate solution, to a final concentration of 8 µM. This incubation occurred at room temperature to monitor the band development, which reached optimal development within 3-5 minutes. The reaction was then stopped by washing the gel in an excess of ice-cold 7.5% acetic acid for 5 minutes at room temperature, followed by rinsing of the gel in water at room temperature for 5 minutes. The gels were then sealed in a plastic sleeve and electronically recorded by scanning the image using an Epson scanner and supplied software (Seiko Epson, Japan). The scanned image was then captured into the GelCompar II (Applied Maths, Belgium) software analysis programme for recording and analysis of the banding profiles. The GelCompar programme also enables one to establish a database of profiles collated with the sample information.

2.3.7 Allele-Specific PCR Analysis of *NAT2*

The 341T>C nucleotide change specifically identifies the *NAT2*5* class of alleles. This polymorphic variation was assessed using the confronting primer PCR technique (CP-PCR) of Hamajima *et al* [251], and is depicted in Fig. 2.4.

Panel (a)



Panel (b)

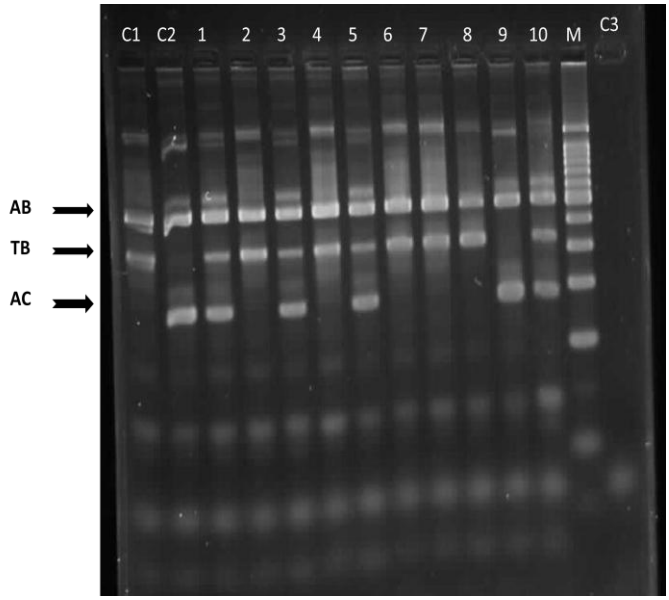


Figure 2.4 Confronting Primer PCR. Two primers (T and C) target the allele specific nucleotide position 341T>C (dotted arrow). Primers A and B facilitate amplification of the 341C (product AC), and/or 341T alleles (product TB) respectively. Primers A and B also generate PCR product AB, which serves as an internal control and monitor of the amplification conditions. Lanes C1, C2, C3 represent the 341TT, 341CC and the negative PCR controls respectively Lane M is the 100bp DNA ladder molecular marker (Promega) Lanes 1 – 10 are test samples.

Two allele-specific, confronting primers, 341T and 341C, were used in a multiplex PCR reaction, along with two other primers, outer-primer A and outer-primer B, (see Table 2.1 for nucleotide sequences) to direct the specific amplification from the 341C-allele or 341T-allele, respectively. Each 25 µl PCR reaction consisted of 20 ng of gDNA, 2.5 µl of a 10x PCR buffer (200 mM Tris (pH 8.4), 500 mM KCl) (Gibco BRL Life Technologies, USA), 0.25 mM of each of dATP, dGTP, dCTP, and dTTP, 2.5 mM $MgCl_2$, 0.25 µM of each of the four primers, and 0.25 units of *Taq* DNA Polymerase I (Gibco BRL Life Technologies, USA). PCR amplification was carried out in the Gene Amp PCR System 9700 (Applied Biosystems, USA).

Table 2.1 Primer sequences for Confronting Primer PCR.

Primer Name	***NAT2* Primer Sequences for allele-specific amplification, (5'to 3')**	**Tm† (°C)**
341T	^{1045}TTCTTCTGCAGGTGACCA**T**^{1063}	57.0
341C	^{1081}ATGTAATTCCTGCCGTCA**G**^{1063}	53.2
Outer-Primer A	^{894}CACATTGTAAGAAGAAACCGG^{914}	52.8
Outer-Primer B	^{1368}AAGATGTTGGAGACGTCTGC^{1349}	54.9

Superscript numbers indicate the nucleotide positions in the gene sequence. The allele-specific nucleotide position (1063) is written as bold and underlined text.
Tm† = Primer melting point temperature as supplied by the manufacturer, Integrated DNA Technologies (www.idtdna.com)

The course of programmed temperature cycling followed an initial denaturation step of 94°C for 5 minutes, subsequently followed by 35 cycles of denaturation at 94°C for 35 seconds, annealing at 58°C for 45 seconds, and polymerisation at 72°C for 60 seconds. This was followed by an extended period of polymerisation period at 72°C for 10 minutes, and then cooling to 4°C for 60 minutes.

PCR sample aliquots (2 µl) were assessed by gel electrophoresis as already described, employing 2% of a high resolution agarose, MetaPhor® agarose (BMA, USA), prepared and electrophoresed with 1x Sodium Borate (5 mM $Na_2B_4O_7.10H_2O$) Solution. In this case the gels were run at a constant 320 volts for 15 minutes, using the GNA 200 gel apparatus (Pharmacia, USA). Under these conditions, the final temperature of the conductance medium was only slightly increased to 38-40 °C, and the resolution of the DNA bands was significantly improved than had been previously found with TBE as conductance medium.

The agarose gels were stained, and photographed as described above.

2.3.8 Primer design and synthesis

Optimal primers sets for the various PCR amplifications were designed using a freeware version of the Primer Premier primer design programme (http://www.PremierBiosoft.com). This programme uses the criteria of Innis and Gelfand [252] in the initial primer sequence selection, and subsequently also analyses each individual primer for hairpin formation, dimer and cross-dimer formations, as well as false priming within the chosen genetic domain. Use of these parameters facilitated the selection of optimised primer sets for each of the various PCR amplification reactions.

The specificity of primer binding to the chosen DNA gene sequence, either *NAT1* or *NAT2*, was analysed by alignment with the sequence databases at the National Center for Biotechnology Information (NCBI) using the BLAST algorithm (Basic Local Alignment Search Tool; http://www.nlm.nih.gov/blast/blast.cgi). All primers were purchased from Integrated DNA Technologies (IDT, Inc. USA) and were supplied as High Pressure Liquid Chromatography (HPLC) purified products to ensure that all primer molecules were of the uniform specified primer length to obviate spurious PCR product formation during amplification.

2.3.9 Analysis of *NAT1* polymorphisms

The *NAT1* genetic sequence (http://www.ncbi.nlm.nih.gov/nuccore/14018355) was analysed using Primer Premier to design the NAT1frwd (5'-^{117}AGGATACCAGTTGG-3') and NAT1rev (5-^{1585}TTCGAAAATTACAACCAAATAACAA-3') primer set. This set of primers was used for the initial amplification of a 1492 bp sequence domain, spanning the major polymorphic sites that have been described for this gene (http://www.louisville.edu/medschool/pharmacology/).

The PCR amplification reaction was carried out in a total volume of 25 µl, and consisted of 15 ng gDNA, 2.5 µl of 10x PCR Buffer (200 mM Tris-HCl (pH 8.4), 500 mM KCl; Gibco BRL Life Technologies, USA), 0.20 mM a dNTP mixture (containing dATP, dGTP, dTTP and dCTP), 0.25 µM of the NAT1frwd and NAT1rev primers, 2.5 mM $MgCl_2$, and 0.25 units of *Taq* DNA polymerase I (Gibco BRL, USA). PCR amplification was carried out in the Perkin Elmer Gene Amp System 9700 (Applied Biosystems, USA).

An aliquot of each sample PCR product was assessed for PCR efficiency and specificity via submarine gel electrophoresis in 1% Seakem® LE agarose (BMA, USA) using 1x Sodium Borate (SB) buffer (5 mM $Na_2B_4O_7.10H_2O$), and a current of 250 volts for 15 minutes in the Omeg mini-gel apparatus (Omeg Scientific, RSA). Subsequently, randomly chosen PCR products were also assayed for *NAT1* sequence-specificity, via restriction endonuclease cleavage (see Fig. 2.1). DNA bands were visualised by GelStar® staining and photographed.

2.3.10 *NAT1* Genotyping via Direct Sequencing

NAT1 genetic polymorphisms have been tabulated in Table 1.1 and 1.2. Twenty one SNPs, as well as several inserted and/or deleted sections within the 3'-UTR of the gene are located within the amplified *NAT1* domain. Eighteen of these SNPs represent the major polymorphisms that have been thus far been delineated in the coding region of this

gene. We therefore decided to assay the amplified *NAT1* domain by direct sequencing to facilitate the cost-effective analysis of these polymorphisms.

2.3.10.1 Preparation of post PCR-reaction for Sequencing

Aliquots of the *NAT1* PCR-reaction product were purified to remove unincorporated dNTPs and primers (NAT1frwd and NAT1rev), via treatment with the ExoSAP-IT® reagent kit (Amersham Pharmacia Biotech., USA). This kit is a propriety formulation of the shrimp alkaline phosphatase (SAP) and exonuclease I enzymes, catalytically active in most PCR reaction buffers. Furthermore, the kit treats post-PCR solutions in which the product size ranges from less than 100 basepairs (bp) to greater than 20 kilo-basepairs (kb). According to the manufacturer's recommendations 2µl of ExoSAP-IT solution reacts with 5 µl of post-PCR reaction volume, degrading all unincorporated primers and dNPTs, with no loss of double-stranded DNA product. Furthermore these volumes can be scaled up proportionally as required.

We mixed 15 µl of each NAT1 PCR product solution with 5 µl of the ice-cold ExoSAP-IT solution, and then incubated this mixture at 37°C for 1 hour, according to the manufacturer's recommendations. This was followed by a further incubation at 80°C for 30 minutes to expedite the complete inactivate of the Exo-SAP-IT enzyme reagent, whereafter the solution can be stored at -20°C until required for sequencing.

2.3.10.2 Sequencing of *NAT1* PCR Products

NAT1 sequencing reactions were facilitated in the ABI PRISM Genetic Analyser model 3130x1 (Applied Biosystems), via an automated cycle sequencing protocol, using the BigDye Terminator ready reaction sequencing kit, version 3.1 (Applied Biosystems), at the Central Analytical Facility (CAF), Stellenbosch University. The kit contains all the required reagents in a premixed solution, to expedite the synthesis, in a single tube, of reaction products labelled with fluorescent dideoxynucleotides (FddNTPs). Essentially the protocol entails a primer extension reaction, in which the FddNTPs are sequentially added to sequence-specific oligomers to form sequenced templates of varying lengths,

terminating with specific dideoxynucleotides. The results of each single reaction tube are then assayed in the genetic analyser, which separates the fragments by electrophoresis, and automatically identifies the separate FddNTPs in the sequenced template, based upon the different emission wavelengths for the fluorescent dye attached to each of the ddNTPs.

One of the disadvantages of this technique evolves around the "spectral deconvolution" of the sequenced results, which limits the amount of nucleotide sequence data that can be read reliably by the software programme to 700 bp. This limitation therefore necessitated sequencing both DNA strands in opposite directions for each sample, to ensure the complete sequence analysis of the 1492 bp *NAT1* PCR product.

Each sample reaction was constituted as follows:- 4 µl of the premixed solution of the sequencing kit, 3 µl of primer (sense or antisense) at a concentration of 1.1 µM, and 3 µl of ExoSAP-IT-treated PCR template solution, and 1.5 µl of dimethyl sulfoxide (DMSO) solution. The latter reagent is added to facilitate the unwinding of potential DNA secondary structure formation in the PCR template. Individual sequencing reactions were carried out in a 96-well plate format, containing both positive and negative control sample reactions randomly spaced among the test samples. All pipetting reactions were executed using the Tecan Evo 150 automated liquid handling workstation (Tecan Group Ltd., Switzerland). The positive control sample template consisted of pGEM plasmid DNA and the M13 primer oligomer, both of which are supplied in the kit. Negative control sample contained no DNA template.

Each thus-prepared 96-well plate is cycle-sequenced in a Gene Amp PCR system 9700 (Applied Biosystems, USA) using a temperature cycling programme, consisting of 1 cycle at 95°C for 5 minutes, followed by 25 cycles of alternating temperatures of 95°C for 45 seconds, 55°C for 4 minutes; the process is terminated with a final incubation at 4°C. These PCR products are then subjected to another clean-up phase, utilising CENTRI-SEP columns (Princeton Separations, USA), to remove buffer salts, unused FddNTPs and oligonucleotide primer via gel exclusion filtration. The use of the CENTRI-SEP 96-well plate facilitates the processing of these reaction samples in keeping with the above multi-well format.

Briefly, CENTRI-SEP Sephadex gel columns are reconstituted and equilibrated by centrifugation at 3000 rpm for 2 minutes in a Heraeus megafuge 11 (Thermo Fischer Scientific, USA). SDS-treated cycle-sequenced PCR products are loaded onto each column in the 96-well plate, and collected by centrifugation at 3000 rpm for 2 minutes as above. Next the reaction products are vacuum dried for 15 minutes at 60°C in a heated vacuum drier (Labconco Corportaion USA), and resuspended in Hi-Di solution (Applied Biosystems, USA). Prior to electrophoresis in an automated sequence analyser the solution is heat-denatured for 2 minutes at 95°C, and then chilled on ice for 5 minutes.

These samples are applied to the ABI-PRISM DNA analyser platform for analysis via a 50 cm long multi-capillary array, employing the Performance Optimum Polymer 7 (POP7), along with the sequencing analysis software (version 5.2) supplied with the instrument (Seq_Z_36_POP7 and Analysis Protocol 3130POP7_BDTv3-KB-Denovo_v5.2). All results were quality controlled by the staff at CAF, and the sequence alignment analysis was carried out by the researcher, using Sequencher, version 4.10.1 (http://www.genecodes.com).

Reagents and Solutions

(All solutions were prepared using reagent grade water prepared via the Synergy Water Purification System (Millipore Corporation, USA))

Acrylamide/bis-Acrylamide	40% Stock Solution (19:1 ratio) purchased from supplier.
Silver Nitrate	5% solution (w/v):- dissolve 5 g of $AgNO_3$ salt into a total volume of 100 ml sterile water.
Cell Lysis Buffer (1L)	320 mM Sucrose (109.54 g/L) 1% Triton X-100 (10 ml of stock/L) 5m M MgCl2 (5 ml of 1M stock) 10m M Tris, pH 7.6 (10 ml of 1M stock)
EDTA (5ooml)	**Ethylene-Diamine-Tetra-Acetate/acetic acid** 0.5 M solution:- dissolve 93.06 g EDTA and 10 g NaOH pellets in 400 ml of sterile water. Adjust to pH 8.0, using 10 M NaOH solution, and top-up to volume. Sterilize by filtration using a 0.2 micron filter unit.
Nuclei Lysis Buffer (1L)	400 mM NaCl 10 mM Tris, pH 8.0 2 mM EDTA, pH 8.0
Proteinase K Solution	**10 mg/ml stock solution** Dissolve 10 mg in 10 ml of sterile water.
Tris	**Tris (Hydroxymethyl) Aminomethane** 1 M Tris, pH 7.6/8.0:- dissolve 121.1 g of crystals (Tris base) in 800 ml of sterile water; adjust to pH using concentrated HCl, and top-up to the desired volume mark. Sterilise by autoclaving.

Sodium Dodecyl Sulphate **SDS**

10% Solution (w/v):- dissolve 10 g of crystals in 100 ml of sterile water.

Sodium Chloride **NaCl**

Saturated NaCl (> 6 M):-dissolve an excess of crystals in 500 ml of sterile water until the salt no longer goes into solution. Sterilise by autoclaving.

Sodium Carbonate **Na_2CO**

6% solution (w/v):- dissolve 60 g of the salt in 1000 ml of sterile water.

Sodium Thiosulphate **$Na_2O_3S_2.5H_2O$**

40 mM solution:- dissolve 496.37 mg of salt per 50 ml of sterile water.

Sodium Borate **$Na_2B_4O_7.10H_2O$**

50 mM solution:- dissolving 19.07 g per 1000 ml of sterile water. This constitutes a 10x stock solution.

TBE Solution **Tris-Borate-EDTA (A 5x stock solution consists of 0.45 M Tris-Borate/0.01 M EDTA per 1000 ml)**

Dissolve 55 g Tris base, and 27.5 g boric acid per 800 ml sterile water. Add 20 ml of 0.5 M EDTA solution, and top-up to 1000 ml.

TE Solution **Tris-EDTA (1x stock solution consists of 10 mM Tris, 2 mM EDTA)**

Add 1 ml of 1 M Tris, pH 8.0, plus 0.4 ml of 0.5 M EDTA to 98.6 ml of sterile water to constitute 100 ml of TE solution.

Equipment

PCR Machines	Gene Amp 2400 and 9700 PCR Systems (Applied Biosystems, USA)
Submarine Gel Apparati	Omeg Mini Gel (Omeg Scientific, Claremont, South Africa GNA 200 Submarine Unit (Amersham Pharmacia Biotech., USA)
Polyacrylamide Gel Apparatus	Mini Protean II Electrophoresis Cell (Bio-Rad Laboratories, Inc. USA)
Power Supply	Consort Electrophoresis Supply Unit, model E844
UV Transluminator	3UV Transilluminator, model LMS-26E (UVP, USA)
Heaters/Incubators	Techne DB 3A, Dri Block (Techne, Cambridge, UK)
Centrifuge	Eppendorf model 5810R (Eppendorf, Hamburg, Germany)
Microfuge	Eppendorf, model 5415D (Eppendorf, Hamburg, Germany)
Shaking Incubator	Hybaid Midi Dual 14 (Hybaid, Ltd. Middlesex, UK)
Scanner	Epson Perfection 1670, model J141A (Seiko Epson, Japan)
Photo Documentation	Kodak Digital Science Electrophoresis Documentation and Analysis 120 System (Eastman Kodak Company, USA)

CHAPTER 3

Population diversity of *NAT2* Alleles in South African Ethnic Groups

3.1 Introduction

The "classic" NAT2 acetylation polymorphism, initially determined via phenotypic analyses, has highlighted several issues pertinent to the understanding of xenobiotic metabolism via acetylation. One such aspect concerns the polymorphic acetylation activity within human populations, variably described as either bimodal [35, 253], or trimodal [37, 38, 222]. Previous acetylation studies have established that the frequency of occurrence of the rapid (F) and slow (S) alleles differs substantially between populations (discussed in Chapter 1), and this would account for the observed variation in overall acetylation. In addition, studies have shown that the frequency of several activity-altering SNPs, characterising the S allele also differ between populations [188, 196, 198]. These variations would subsequently impact upon the frequency of slow acetylators between different populations. Furthermore, it has been well established that the NAT2 enzyme can abrogate the activity of isoniazid (INH) [35]. These factors, along with the increasing global TB/HIV pandemic affecting some counties more than others, have facilitated a renewed interest into the effects of acetylation on the epidemiology of TB and HIV [211, 243, 254-256].

Globally, the greatest burden of TB disease occurs in 22 countries (Fig. 3.1), predominantly located in Sub-Saharan Africa and Southeast Asia (WHO Factsheet No. 104, 2010). A few studies have investigated acetylator polymorphisms in these regions [167, 191, 194, 234, 257, 258], but the impact of this variation on drug metabolism and TB epidemiology is largely unknown. Therefore it may be important to examine the *NAT2* polymorphisms to assess whether this impacts on anti-TB therapy. South Africa is ranked fifth amongst the countries with the highest burden of TB disease, and due to the limited information available in the literature on the NAT2 acetylation polymorphism in South Africa, we decided to investigate the acetylation genotypes within local ethnic populations of the Western Cape region of South Africa.

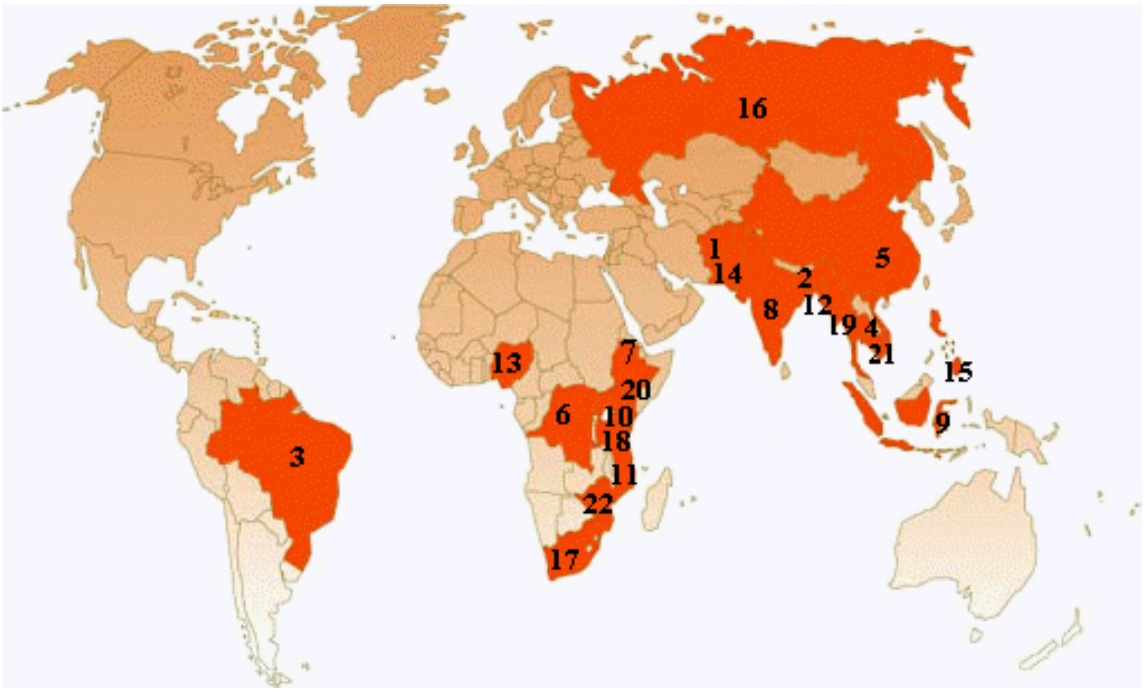


Figure 3.1 A map showing the 22 countries with the highest burden of TB disease globally. India (8), China (5), and Indonesia (9) are ranked first, second and third respectively, whilst South Africa (17) is ranked fifth. Numbers indicate the alphabetical listing for these countries:- Afghanistan (1), Bangladesh (2), Brazil (3), Cambodia (4), China (5), Democratic Republic of Congo (6), Ethiopia (7), India (8), Indonesia (9), Kenya (10), Mozambique (11), Myanmar (12) Nigeria (13), Pakistan (14), Philippines (15), Russian Federation (16) South Africa (17), United Republic of Tanzania (18), Thailand (19), Uganda (20), Viet Nam (21) and Zimbabwe (22). (http://www.stoptb.org/countries/tbdata.asp)

3.2 Study cohorts

The predominant ethnicities in the Western Cape Province are Caucasian (18.4%), Black African (Xhosa) (30.1%), and the group officially classified as South African Coloured (SAC) (50.2%) (http://www.statssa.gov.za/publications/P0301). The incidence of TB in this region has been recorded at greater than 1000 per 100 000 of the population [259]. Furthermore, TB surveillance studies of an urban community in Cape Town, namely, residents of the Ravensmead-Uitsig housing estate (R/U), have shown a

consistent increase in the notification rate for smear-positive TB since 1994 [260, 261], despite having a low prevalence of HIV in the area [262].

The R/U housing estate covers an area of 2.4 km^2 and has a population of 36 334 individuals, 98% of whom are of the SAC ethnic population (Statistics South Africa: Census 2001 http://www.statssa.gov.za). The high burden of TB disease in this area, and the singularity of ethnicity [245, 263], makes this area an ideal study area and cohort for investigations of TB disease. These projects have been approved by the Institutional Research Ethics Committees of Stellenbosch University, and Tygerberg Hospital.

In order to develop a greater understanding of the role of the NAT2 acetylation polymorphism, we randomly recruited non-TB diseased (no active TB), adult individuals from the ethnic groups within the Cape Town Metropolitan region. This collection consisted of 60 Caucasians, 55 Xhosa and 108 SAC individuals. The Caucasian, Xhosa and SAC individuals self-reported their respective ethnicities. The SAC individuals were recruited from the R/U community, and consisted of 46 males and 62 females, with a median age of 34 years (range 18 to 64). The Caucasian group consisted of 24 males and 36 females, with a median age of 74 years (range 39 to 86). These individuals were recruited from the Ophthalmology general clinic at Tygeberg Hospital as the control group for a study investigating age-related macular degeneration (AMD). None of these individuals reported any systemic disease symptoms. The Black African (Xhosa) individuals were randomly selected from healthy volunteers provided from a cancer research study, done by the Medical Research Council (MRC) of South Africa.

3.3 Results and Discussion

3.3.1 *NAT2* SNP and Allele frequencies within cohorts

The NAT2 acetylation polymorphism was characterised by analysis of the 191G>A, 282C>T, 341T>C, 481C>T, 590G>A, 803A>G and 857G>A SNPs, in the individuals from the local cohorts. All SNPs were tested for Hardy Weinberg Equilibrium

(HWE), using HaploView version 4.2 (http://www.broadinstitute.org/mpg/haploview). In both the Caucasian and Xhosa cohorts all of the genotypes were in HWE, indicating panmixia within these specific cohorts. In the case of the SAC only the 803A>G genotype was not in HWE ($p = 0.0318$), indicating a departure from the panmictic status with respect to this polymorphism. Since a similar disequilibrium was not observed in either of the Caucasian or Xhosa cohorts for this particular SNP, it is unlikely that this departure from HWE in the SAC cohort is due to error or bias in the experimental assay.

The predominant or major *NAT2* allele at each of the respective SNP positions was classified as the wild-type (WT) allele, whilst the polymorphic variant, is classified as the minor allele (MA) (Table 3.1). The *NAT2* allele frequency observed in the SAC and Xhosa cohorts appear very similar, and there are no significant differences between these two groups. The Caucasian cohort on the other hand appears to constitute a different group, in which the MA frequencies for 4 of the 7 SNPs assayed were generally higher than observed in either the SAC or Xhosa groups. In addition, the 191A allele was not found in the local Caucasian population.

The relevance of the various SNPs depends upon the phenotypic effect conferred by that polymorphism. In the case of the 282C>T and 481C>T polymorphisms, the acetylation activity of the enzyme is supposedly not impaired, since these SNPs encode the synonymous amino acids tyrosine and leucine at positions 94 and 161 respectively. However, silent or synonymous polymorphisms in the Multidrug Resistance 1 (MDR1) gene have recently been shown to affect the substrate specificity of the enzyme, without concomitant changes in the expression levels of MDR1 mRNA or protein [264, 265]. In this context, the significance of the above observations with respect to the 282C>T and 481C>T *NAT2* SNPs should be further investigated, to assess their affect on INH metabolism, particularly in these local ethnic groups, given the observed allele frequency for these two SNPs. However, this was not investigated at this time.

A Pearson Chi square (χ^2) analysis of the allele frequency for these particular SNPs indicates a significant difference between the three cohorts (Table 3.1), which was also found when comparing the Caucasian and SAC groups only. In the latter case the Chi

square test rendered *P*-values of 0.0178 for the 282C>T polymorphism, and 0.0190 for the 481C>T polymorphism, respectively.

Similarly, the 803A>G polymorphism, which also confers rapid acetylation (RA) status, changes lysine to arginine at position 268. In the case of this SNP a comparison between the Caucasian and SAC cohorts revealed an allele frequency difference of borderline significance only (*P*-value of 0.0408).

Table 3.1 Prevalence of the *NAT2* alleles in local Ethnic groups

Prevalence of *NAT2* Alleles					***P* - Value[§]**		
***NAT2* allele**		**Caucasian (n=120)**	**Xhosa (n=110)**	**SAC (n=216)**	**Caucasian / Xhosa / SAC[Δ]**	**Caucasian / Xhosa[Δ]**	**Caucasian / SAC[Δ]**
191	**G**[†]	100	0.98	0.95	0.0358	0.1379	0.0172
	A	0	0.02	0.05			
282	**C**[†]	0.67	0.70	0.78	0.0375	0.6942	0.0178
	T	0.33	0.30	0.22			
341	**T**[†]	0.62	0.73	0.70	0.1733	0.0749	0.1564
	C	0.38	0.27	0.30			
481	**C**[†]	0.64	0.78	0.76	0.0250	0.0195	0.0190
	T	0.36	0.22	0.24			
590	**G**[†]	0.73	0.86	0.87	0.0025	0.0151	0.0011
	A	0.27	0.14	0.13			
803	**A**[†]	0.58	0.46	0.46	0.0874	0.0678	0.0408
	G	0.42	0.54	0.54			
857	**G**[†]	0.97	0.97	0.96	0.8010	0.7892	0.7161
	A	0.03	0.03	0.04			

[†] = Wild type or major allele.
[§] = *P*-values determined by Pearson Chi-square (χ^2) analysis; a *P*-value of less than 0.05 indicates statistical significance.
[Δ] Statistical comparisons of the *NAT2* alleles between the various ethnic groups

The 191G>A, 341T>C, 590G>A and 857G>A polymorphisms reduce the acetylation activity of the NAT2 enzyme to varying degrees [168, 182]. In this respect the 341T>C and 857G>A alleles did not show any substantial differences between the cohort groups investigated. The most significant differences were however observed for the 191G>A, and 590G>A alleles. In the SAC population group, the 191A allele was 2.5-fold more prevalent than in the Xhosa; however, this difference was not found to be statistically significant by Pearson Chi square (χ^2) analysis (*P*-value = 0.207).

The most significant differences between the three groups was observed for the *NAT2* 590G>A polymorphism, which showed an even greater significance in the comparison between the Caucasian and SAC cohorts (*P* = 0.0011). Furthermore, the prevalence of the 590A allele varies between 13% and 27% in the three cohorts investigated, which indicates that this SNP would make a relatively significant contribution to the variation in the inherent acetylator activity observed in these cohorts. The significance of the observed allelic differences between the two most divergent groups (Caucasians and SAC) is provocative, but only translates into a relatively low probability of having an impact on the acetylation differences between these groups, given the relatively small sample size (only a 168 individuals). Consequently, assuming a power of 80% (probability to detect an effect), it would be necessary to investigate a much larger cohort of individuals (approximately 1100) in order to determine whether these allelic differences could have a significant effect on the acetylation differences between these groups.

A comparison of the *NAT2* genotypes defined by these seven SNPs corroborates the results observed for their *NAT2* alleles (Table 3.2). In this context the 282C>T and 481C>T genotypes are found to be significantly different between the Caucasian and SAC cohorts, whilst the 803A>G genotype is not (Table 3.2). Furthermore, the 191GG and 191GA genotypes were significantly different between the Xhosa and SAC cohorts, thereby indicating that a variation in the slow acetylation status between these cohorts could be attributed to this particular polymorphism. However, the impact of this variation upon the overall acetylation metabolism warrants further investigation, since the allele and genotype frequencies for the 191G>A SNP is less than 10% in these cohorts.

The genotypes conferred by the 590G>A polymorphism again showed the most significant variation between the cohorts. This indicates that this SNP would confer the greatest impact on the acetylation status between the different cohorts (*P*-value = 0.008), with an even greater significant difference between the Caucasian and SAC cohorts as assessed by the Fischer's exact test (*P*-value = 0.002). Given the observed prevalences for the 590G>A polymorphism between these cohorts, it is clear that this particular SNP would exert the greatest influence upon the overall phenotypic acetylation status. Therefore the variation in the minor allele frequencies and genotype for this particular SNP may have a significant bearing upon the acetylation status in different ethnic groups.

Table 3.2 Comparison of the *NAT2* genotypes in three local ethnic groups.

***NAT2* SNP**		***NAT2* Genotype Frequencies (%)**			***P* – Value**		
		Caucasian (n = 60)	**Xhosa (n = 55)**	**SAC (n = 108)**	**Caucasian / Xhosa / SAC[§]**	**Caucasian / Xhosa[‡]**	**Caucasian / SAC[‡]**
191	**GG**	100	96.4	90.7	0.033	0.228	0.016
	GA	0	3.4	9.3			
	AA	0	0	0			
282	**CC**	43.3	45.5	57.4	0.089	0.778	0.02
	CT	46.7	49.1	40.7			
	TT	10	5.4	1.9			
341	**TT**	38.3	49.1	50.9	0.204	0.0922	0.19
	TC	46.7	47.3	38.0			
	CC	15	3.6	11.1			
481	**CC**	43.3	58.2	59.3	0.053	0.021	0.023
	CT	41.6	40.0	34.3			
	TT	15	1.8	6.4			
590	**GG**	55	74.6	74.1	0.008	0.016	0.002
	GA	35	23.6	25.0			
	AA	10	1.8	0.9			
803	**AA**	30	21.8	15.7	0.119	0.086	0.053
	AG	55	49.1	61.1			
	GG	15	29.1	23.2			
857	**GG**	93.3	94.5	91.7	0.794	0.100	1.000
	GA	6.7	5.5	8.3			
	AA	0	0	0			

‡ = Fischer's exact Test, two-tailed analysis.
§ = Pearson chi-square (χ^2) analysis.

3.3.2 *NAT2* Genotypes and Acetylation within local cohorts

Individual *NAT2* genotypes were reconstructed in accordance with a common convention in the literature, in which rapid acetylation (RA) is represented by two wild-type alleles (F), characterised by a combination of any 2 of the haplotype combinations delineated by the following *NAT2* alleles:- *NAT2*4, NAT2*11, NAT2*12* and *NAT2*13.* The reference allele is defined as *NAT2*4*, and confers the WT rapid acetylation capacity. The *NAT2*11*, *NAT2*12* and *NAT2*13* alleles and genotypes also confer the RA status, but are defined by polymorphisms at positions 481, 803 and 282 respectively. Similarly, slow acetylation (SA) status is conferred by two S alleles, characterised by the *NAT2*5*, *NAT2*6*, *NAT2*7* and *NAT2*14* haplotypes and alleles, (see table 1.4 in Chapter 1). The intermediate acetylation (IA) status and genotype by comparison, is represented by the heterozygous FS allele assignment.

The final *NAT2* genotypes were deduced from the sum of the polymorphisms observed for the 7 *NAT2* SNPs, according to Agúndez *et al* (2008) [266]. This classification scores the polymorphic variation at each individual SNP as zero (0) for no polymorphism (WT allele), 1 for SNP heterozygosity, and 2 for SNP homozygosity. Secondly, the SNPs are arranged sequentially, in accordance with their occurrence in the coding region of the *NAT2* gene, namely 191G>A, 282C>T, 341T>C, 481C>T, 590G>A, 803A>G and 857G>A. Thus the *NAT2*4/2*4* genotype, in which no polymorphisms occur at any of the above SNP positions, is assigned the "0000000" classification for the genotype (combined diplotype) (Table 3.3).

According to this classification, some of the observed "diplotypes" describe various deduced genotypes, differing in their combinatorial assignment for the 7 *NAT2* SNPs. This ambiguity is a consequence of the genotyping methods currently used, in that both the paternal and maternal alleles are analysed simultaneously for each individual sample [83, 267-270]. In this respect, the PCR-RFLP experimental technique used during this project is similarly unable to resolve the phasing of the separate loci with respect to each other, as well as with respect to the separate parental allele(s), during such multi-locus analyses. However, in our analysis these ambiguous genotypes did not result in a concomitant change in the acetylation status as conferred by the classification assignments. Nevertheless, the reliability of this 7-SNP assay and PCR-RFLP

methodology has been confirmed in several publications [266, 269, 271, 272] as well as via computational analyses [266, 270, 273, 274].

Table 3.3 *NAT2* genotypes observed in the local ethnic groups.

Observed Diplotype	Deduced Genotype	Frequency (%)	Alternative Genotype
Caucasian Cohort (n = 60)			
Rapid Acetylators (n = 8)			
0000000	*2*4/2*4*	50	-
0000010	*2*4/2*12A*	25	-
0100000	*2*4/2*13*	25	-
Intermediate Acetylators (n = 22)			
0011010	*2*4/2*5B*	40.9	-
0100100	*2*4/2*6A*	18.2	*2*13/2*6B*
0100110	*2*4/2*6C*	18.2	*2*12A/2*6A*
0111010	*2*4/2*5G*	13.6	*2*13/2*5B* or *2*12B/2*5A*
0010010	*2*4/2*5C*	4.6	*2*12A/2*5D*
0000100	*2*4/2*6B*	4.6	-
Slow Acetylators (n= 30)			
0111110	*2*5B/2*6A*	36.7	*2*5G/2*6B*
0022020	*2*5B/2*5B*	30.0	-
0200200	*2*6A/2*6A*	16.7	-
0111011	*2*5B/2*7B*	10.0	*2*5G/2*7A*
0100200	*2*6A/2*6B*	3.0	-
0200101	*2*6A/2*7B*	3.0	-
Xhosa Cohort (n = 55)			
Rapid Acetylators (n = 15)			
0000020	*2*12A/2*12A*	33.3	-
0100010	*2*4/2*12B*	20.0	*2*12A/2*13*
0000000	*2*4/2*4*	13.3	-
0000010	*2*4/2*12A*	6.7	-
0100000	*2*4/2*13*	6.7	-
0100020	*2*12A/2*12B*	6.7	-
0001010	*2*12A/2*11*	6.7	-
0200000	*2*13/2*13*	6.7	-
Intermediate Acetylators (n = 29)			
0111010	*2*4/2*5G*	24.1	*2*13/2*5B or 2*12B/2*5A*
0011020	*2*12A/2*5B*	17.2	-
0100110	*2*4/2*6C*	10.4	*2*12A/2*6A or 2*12B/2*6B*
0011010	*2*4/2*5B*	6.9	*2*12A/2*5A*
0100100	*2*4/2*6A*	6.9	*2*13/2*6B*
0000100	*2*4/2*6B*	6.9	-
0010020	*2*12A/2*5C*	6.9	-
0110010	*2*13/2*5C*	6.9	*2*12B/2*5D*
0010000	*2*4/2*5D*	3.5	-
0000110	*2*4/2*6F*	3.5	*2*12A/2*6B*
0101011	*2*12C/2*7B*	3.5	-
0200100	*2*13/2*6A*	3.5	-
Slow Acetylators (n = 11)			
0111110	*2*5B/2*6A*	27.3	2*5G/2*6B
0021020	*2*5B/2*5C*	18.2	-
1111010	*2*5B/2*14B*	18.2	*2*5G/2*14A or 2*5A/2*14G*
0022010	*2*5B/2*5A*	9.1	-
0111011	*2*5B/2*7B*	9.1	*2*5G/2*7A*

0100200	*2*6A/2*6B*	9.1	-
0200111	*2*6C/2*7B*	9.1	-
SAC Cohort (n = 108)			
Rapid Acetylators (n = 24)			
0000010	*2*4/2*12A*	41.7	-
0000020	*2*12A/2*12A*	25.0	-
0000000	*2*4/2*4*	12.5	-
0100010	*2*4/2*12B*	12.5	*2*13/2*12A*
0001010	*2*4/2*12C*	4.2	*2*11/2*12A*
0100000	*2*4/2*13*	4.2	-
Intermediate Acetylators (n = 55)			
0100110	*2*12A/2*6A*	23.6	*2*4/2*6C*
0011010	*2*4/2*5B*	20.0	*2*12A/2*5A*
0011020	*2*12A/2*5B*	9.1	-
0010010	*2*4/2*5C*	5.5	*2*12A/2*5D*
0111010	*2*4/2*5G*	3.6	*2*13/2*5B*
0100100	*2*4/2*6A*	3.6	*2*13/2*6B*
0100001	*2*4/2*7B*	3.6	*2*13/2*7A*
0010020	*2*12A/2*5C*	3.6	
0100011	*2*12A/2*7B*	3.6	-
1000010	*2*12A/2*14A*	3.6	*2*4/2*14E*
1100010	*2*12A/2*14B*	5.5	-
0110010	*2*12B/2*5D*	3.6	*2*13/2*5C*
0010000	*2*4/2*5D*	1.8	-
1000000	*2*4/2*14A*	1.8	-
1100000	*2*4/2*14B*	1.8	*2*13/2*14A*
0000110	*2*12A/2*6B*	1.8	*2*4/2*6F*
0110000	*2*13/2*5D*	1.8	-
0200100	*2*13/2*6A*	1.8	-
Slow Acetylators (n = 29)			
0022020	*2*5B/2*5B*	24.1	-
0111110	*2*5B/2*6A*	20.7	*2*5G/2*6B*
0021020	*2*5B/2*5C*	17.2	-
1011010	*2*5B/2*14A*	6.9	-
0111011	*2*5B/2*7B*	6.9	*2*5G/2*7A*
0011110	*2*5B/2*6B*	3.5	*2*5C/2*6E or 2*5A/2*6F*
0111100	*2*5A/2*6A*	3.5	*2*11A/2*5J or 2*5K/2*6E*
0010011	*2*5C/2*7A*	3.5	-
0111001	*2*5A/2*7B*	3.5	-
0100200	*2*6A/2*6B*	3.5	-
0200101	*2*6A/2*7B*	3.5	-
1100100	*2*6B/2*14B*	3.5	*2*6A/2*14A*

NAT2 genotypes are represented as a 7-digit combinatorial score for positions 191, 282, 341, 481, 590, 803, and 857 respectively.

0 = no polymorphism, 1 = heterozygous polymorphism, and 2 = homozygous "mutant" polymorphism for each of the polymorphic loci analysed

The overall *NAT2* acetylation status, as defined by these deduced genotypes is summarised in figure 3.2. The Xhosa and SAC cohorts exhibit similar frequencies for the RA, IA and SA classes of trimodal acetylation, with the IA status predominating in each of these groups. In a study of SAC TB patients, Parkin *et al* [170] reported frequencies of 23%, 47% and 30% for RA, IA and SA, respectively. Similar frequencies were found for SAC individuals in this study, namely 22.2% (RA), 50.9% (IA) and 26.9% (SA).

In the Caucasian group, slow acetylators (SA) predominate by a factor of two, when compared to the Xhosa and SAC cohorts. In contrast, fast (RA) acetylators are approximately twice as prevalent in the Xhosa and SAC cohorts compared to Caucasians. These results suggest that the overall rate of metabolism of xenobiotics via NAT2 acetylation is similar in Xhosa and SAC individuals, and differs significantly from that observed for Caucasians. This is supported by the frequency distribution for the rapid (R) and slow (S) alleles in these three cohorts (Fig. 3.3). Therefore, anti-TB regimens containing INH and standardised for Caucasian individuals may not provide sufficient treatment efficacy in the Xhosa and SAC population groups, given the difference in rapid and intermediate acetylators observed in these groups.

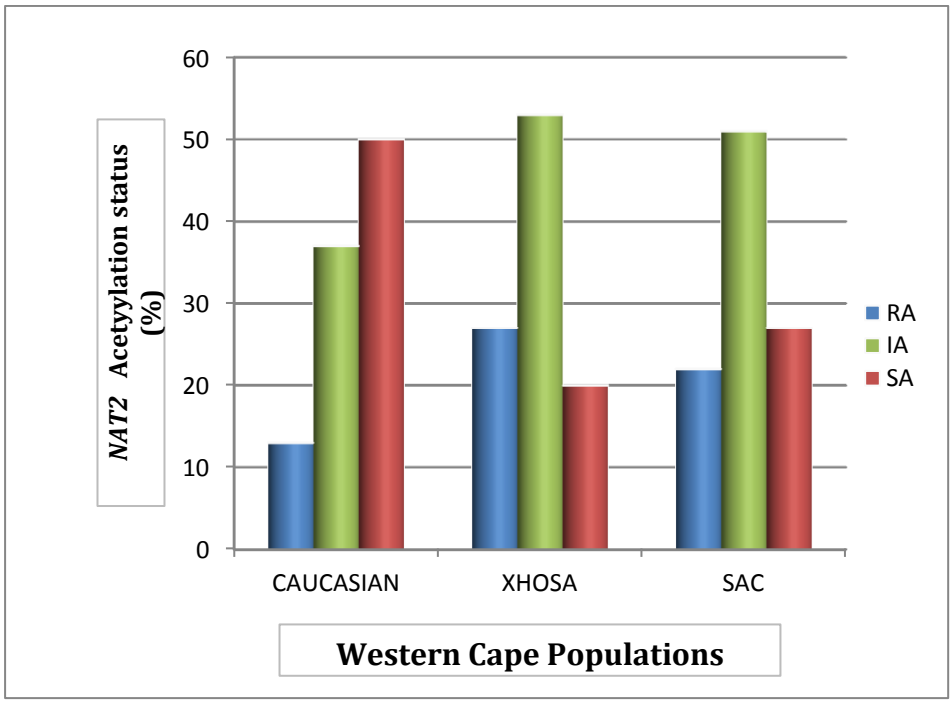


Figure 3.2 *NAT2* acetylation status in ethnic groups of the Western Cape.

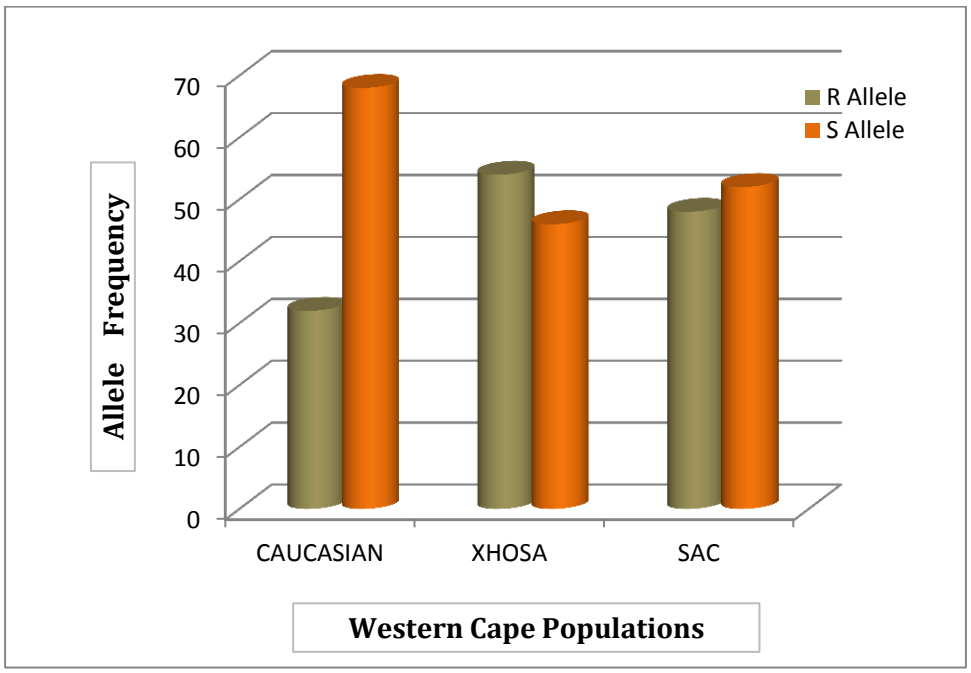


Figure 3.3 Distribution of the wild type (R) and polymorphic (S) *NAT2* alleles in local cohorts.

A comparison of the *NAT2* haplotypes, as defined by the seven SNPs investigated is depicted in figure 3.4. Frequency distributions for the *2*4*, and *2*12* haplotypes are similar in the Xhosa and SAC groups, but are significantly different in the Caucasians. On the other hand, the frequency distribution of the *2*13* haplotype is not significantly different in these three cohorts. Hence the difference in the rapid acetylation mechanism between Xhosa and SAC individuals on the one hand and Caucasians on the other, would largely be defined by the *NAT2*4* and *2*12* haplotypes.

The *NAT2*5* haplotype is the predominating slow acetylator genotype in all three study groups, occurring at a comparable frequency of 55 - 61%. Secondly, the *2*6* haplotype is 1.5x more frequent in Caucasians than in the other two ethnic groups, and together with the *2*5* haplotype accounts for 95% of the slow acetylation genotype and phenotype in Caucasians. In contrast, the *2*5* and *2*6* haplotypes

contribute only 83% towards the slow acetylation genotype in the SAC ethnic group.

The *2*7* and *2*14* haplotypes are more prevalent in SAC individuals, with a combined frequency of approximately 17 %. By comparison, the *2*14* haplotype was not found in Caucasians, and occurred at a frequency of only 3.9% in the Xhosa group. Therefore the prevalence of these haplotypes (*2*7* and *2*14*) could confer a greater slow acetylation effect in SAC individuals, than Xhosas or Caucasians.

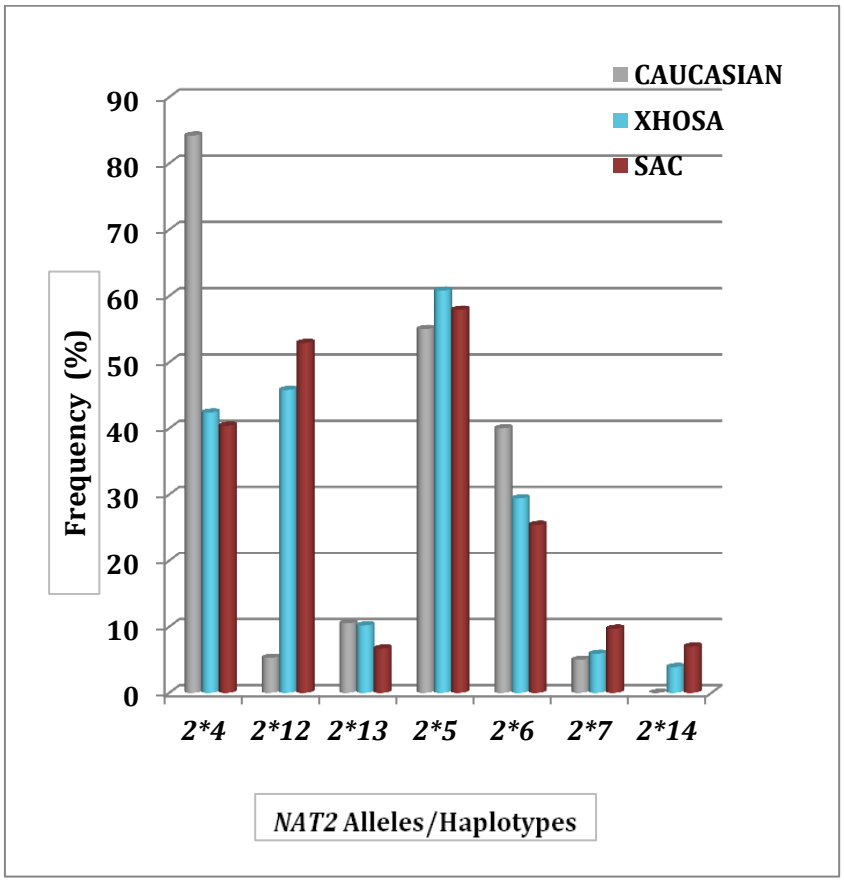


Figure 3.4 Comparison of *NAT2* alleles and haplotypes in local cohorts.

3.4 Conclusion

INH is the primary drug used in anti-TB treatment regimens, and it is metabolised by NAT2, a human liver enzyme. This inactivation of INH would in turn impact upon the treatment efficacy and consequently also the control of TB-disease. We investigated the NAT2 acetylation status in the three major ethnic groups in a TB endemic area of South Africa (Western Cape region). This region has a recorded TB incidence of greater than 1000 per 100 000 of the population [259]. A panel of *NAT2* SNPs, having a well-characterised genotype-phenotype concordance, were used to assess the overall metabolic activity attributable to rapid (RA), Intermediate (IA), or Slow (SA) acetylation in these ethnic groups (http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/).

Our results indicate that there is an overall difference in the relative proportions of rapid (RA), intermediate (IA) and slow acetylators (SA) between the three study cohorts. The significance of this observation warrants further investigation. The SA were found to be more prevalent in Caucasians, which is comparable with similar results reported for Caucasians in Europe [62, 192, 196] and North America [164]. In Xhosa and SAC individuals the IA predominated, occurring at relatively comparable frequencies within these two populations. In both populations the RA and SA occurred at similarly low frequencies, with the SA being somewhat more prevalent in SAC compared to Xhosa individuals. A recent study by Luca *et al* (2008) [275] has shown that SA and IA predominate in a geographic belt from Eastern Africa to North-western Europe, whereas RA were more prevalent in Western Africa and Northern Eurasia. In this context, the acetylation trend amongst the SAC would appear to correspond more to the East African population, whereas the Xhosa group appears similar to Western African populations.

This result can be explained by the prevailing evidence which suggests that Bantu-speaking people migrated out of West Africa, initially in an easterly direction, followed by a strong southerly migration. These agriculturalists settled as far south as the Fish and Kei rivers in South Africa, where intermingling with the Khoisan occurred (hunter-gatherer/pastoralists) [276]. Thus the Xhosa, as can be detected in their click-language dialect, are themselves admixed with the Khoisan. The Bantu origins of the Xhosa would

therefore contribute a strong RA input, whilst the admixture with the Khoisan would contribute the SA status, essentially leading to a dominance of IA.

Whilst the significance of these differences is still speculative, it has been proposed that acetylation differences between diverse population groups arose as a result of adaptations to changes in the environment and diet that occurred during the evolution of humans from nomadic hunter-gatherers (such as the Khoisan) to sedentary farmers or agriculturalists (such as the Bantu) [275, 277, 278]. Furthermore, these authors found that RA is more prevalent in hunter-gatherer populations, such as the Pygmies and the Khoisan, whilst SA is more prevalent in populations practising an agricultural lifestyle and existence. The current SAC population is also admixed, with genetic input from Khoisan (32-43%), Bantu (20-36%), Caucasian (21-28%) and Asian (9-11%) groups respectively [245]. One may therefore have expected the Khoisan genetic input into the SAC to have biased the *NAT2* genotype towards RA. However, the greater contribution of the Bantu and Caucasian ancestors to the genetic heritage of the SAC clearly contributes strongly to the SA status. Thus the dominance of the IA in the SAC is not surprising, given the strong genetic contributions from these respective ancestral groups.

This result suggests that the standardised INH dosage, given in an anti-TB treatment regimen to Caucasian individuals, would be subject to a lower rate of acetylation, and hence inactivation, compared to the Xhosa and SAC individuals in these local populations. Therefore a better treatment efficacy may be attained in Caucasians, than in Xhosa or SAC individuals. Given the relatively high numbers of intermediate and rapid acetylators in the Xhosa and SAC groups, compared to Caucasians, it is likely that such a standardised INH dosage could be less effective in these individuals, due to the increased acetylation of INH. Therefore, it is likely that similar acetylation pharmacodynamics exist in these groups, which may differ from that for Caucasians, and it may be expedient to monitor individuals from these ethnic groups more closely during anti-TB therapy.

The frequency distribution for the separate *NAT2* alleles and haplotypes in the study cohorts (Fig. 3.4) depicts a degree of complexity evident in the level of rapid and slow acetylation conferred by the *2*12*, *2*6*, *2*7*, and *2*14* haplotypes. *NAT2*12* encodes the amino acid change lysine to arginine at position 268 in the enzyme, thereby conferring the

rapid acetylation phenotype. Earlier studies assessed NAT2 acetylation-capacity by using caffeine as the probe drug in human subjects [43, 174], and since caffeine is initially and primarily metabolised by CYP1A2 [279-281], it may therefore not undergo a similar acetylation rate to INH. In addition, CYP1A2 is subject to both inducible and constitutive expression which would have an additional confounding effect on the downstream metabolism of caffeine by NAT2. Furthermore, a report by Kawamura *et al* (2005) [97] indicates that NAT2 may have a greater substrate affinity for INH, based on the greater acetylation activity of INH compared to caffeine. In this context, it is therefore not yet clear what the functional effect of this particular SNP is on enzyme activity, since supporting functional experimental data is outstanding [140]. Therefore given the differences observed in the study cohorts, it would be expedient to investigate whether the *NAT2*12* rapid phenotype is equal to or greater than that conferred by *NAT2*4.*

In the case of the slow acetylation haplotypes there is no difference between the cohorts with respect to the rate of slow acetylation attributable to the *2*5* haplotype. A similar observation is seen for the combined effects of the *2*6*, *2*7* and *2*14* haplotypes (fig. 3.4), which collectively contribute between 39 – 45% towards slow acetylation. However, in vitro studies have demonstrated that the proteins expressed by the 2*7 and 2*14 variants were the least stable (compared to NAT2*4) [175], and hence (slow) acetylation capacity is significantly reduced when compared to the 2*5 and 2*6 slow acetylators [182]. Therefore, the prevalence of these haplotypes in the few SAC individuals identified in this study, may reflect even lower acetylation capacity compared to Xhosa and Caucasians. SAC individuals harbouring homozygous or heterozygous combinations of 2*7 and or 2*14 haplotypes may also experience low levels of acetylation, and may therefore exhibit a better treatment response, due to the reduced acetylation of INH.

Altered rates of acetylation have also been shown in *in vivo* studies using caffeine as phenotypic probe substrate [174], where it was demonstrated that some heterozygous allele combinations have a lower acetylation activity when compared to homozygous allele combinations. This would become more evident in ethnic groups exhibiting a wider spectrum of particular genetic haplotypes, since a greater number of allelic permutations could arise and thereby constitute a larger spectrum of phenotypes within the group.

Our analysis of NAT2 genotypes in three local cohorts in the Western Cape region of South Africa has indicated that the spectrum of NAT2 haplotypes differs between these groups, with some haplotypes being absent from a particular group, such as *2*14* which was not found in Caucasians. In addition, some haplotypes are more prevalent in one particular group, such as *2*4* in Caucasians, and *2*7* and *2*14* in SAC, whilst others, such as *2*12*, are equally prevalent in SAC and Xhosa. In a similar study by Parkin *et al* [63], greater frequency for the *2*12* allele was also observed in SAC TB patients. In comparing *NAT2* haplotypes between African and non-African populations, Sabbagh *et al* (2011) [278] reported a higher frequency for the *2*12* and *2*13* haplotypes in African populations, whereas, the *2*4* haplotype was more prevalent in non-African populations. We observed a similarly high prevalence for the *2*12* and *2*13* haplotypes (alleles) in the SAC and Xhosa groups, which confirms this observation for African populations.

In past genetic studies the importance of low frequency haplotypes, such as *2*7* and *2*14*, may have been overlooked, since it was believed that these haplotypes would not have a major impact on the manifestation of the disease phenotype. However, recent developments in genetic analyses have enabled genome-wide association studies (GWAS) to investigate polymorphisms more comprehensively, on a genomic, rather that a genetic scale [282-284]. The results from these studies indicate that common (more prevalent) polymorphisms only constitute a relatively small contribution to the overall disease phenotype. Low frequency polymorphisms on the other hand, having a functional effect may constitute better candidates for understanding variations in phenotype. In this regard it may be expedient to re-evaluate *NAT2* SNPs with respect to their phenotypic acetylation of INH.

CHAPTER 4

The analysis of *NAT2* Acetylation polymorphisms in TB patients

4.1 Introduction

TB surveillance studies conducted in the Ravensmead-Uitsig housing estate (R/U) - an urban community in Cape Town - have indicated a consistent increase in the notification rate for smear-positive TB since 1994 [260]. Approximately 60% of patients from this area are new TB cases [285]. Similarly, a high prevalence of clinical TB has also been reported in patients from this area who have previously been treated and ostensibly cured of their disease [259], which would indicate that treatment efficacy within this community is problematic, thereby arguably worsening the TB epidemic. A report by Kritzinger *et al* [262], noted an annual rate of TB infection in this community of greater than 3.5%, which constitutes one of the highest figures reported globally for TB disease.

The R/U housing estate covers an area of 2.4 square km, and has a population of 36 334 individuals, 98% of whom are of the South African Coloured (SAC) ethnic population group (Statistics South Africa: Census 2001 http://www.statssa.gov.za). Given the high burden of disease in this area and the singular ethnicity [263], together with a relatively uniform set of contributory environmental factors, this community is an ideal study cohort for investigating the epidemiology of TB. In this context, it is significant to note that the prevalence of HIV infection in the SAC community is lower than for the whole of the Western Cape province, which co-incidentally has the lowest HIV prevalence in South Africa [286]. Subsequently, several longitudinal research projects have been conducted by the Faculty of Health Sciences of Stellenbosch University, along with the assistance and active participation of this community since 1992. All projects have been approved by the Human Research Ethics Committee of Stellenbosch University.

The South African Coloured (SAC) population is officially recognised as a characteristic grouping of individuals, whose genetic ancestry devolves from Africa, Europe, Asia and Madagascar [244, 287-289]. In the ensuing generations, since the colonization of the southern tip of Africa by the Dutch in 1652, the genetic admixture within this group has coalesced into a relatively homogenous ensemble [290], in which the major contributing antecedents are Khoi and San (32-43%), Bantu (20-36%), European (21-28%), and Asian (9-11%) [245]. The Khoi and San are the considered indigenous groups to the area, and

provided the genetic background assimilating the relatively new genetic inputs from Europe, Africa (Bantu) and Asia. Hence the SAC are endowed with a genetic similitude, which is also evident in the distinctive frequency distribution reported for the human leukocyte antigen (HLA) class of alleles. In the case of the HLA-class II alleles, the SAC group has a frequency distribution unlike any of the other southern African population groupings [291], providing further support for the consolidation of the SAC group-identity within the Western Cape.

TB patients from the R/U community were enrolled in a longitudinal prospective study project, the Surrogate Marker (SM) study, and monitored for a period of twenty-four months, from their first diagnosis of clinical disease (enrolment) [292, 293]. A major aim of the project was to collect, analyse, and evaluate a range of patient data to identify a biomarker (or biomarkers) which could serve as a surrogate to predict the outcome of treatment for active TB. Host derived factors, such as malnutrition, BCG vaccination, HIV status, age and gender, have been shown to influence the advent of the clinical disease [294, 295]. Furthermore there is epidemiological evidence to suggest that host derived factors may also be implicated in the individual's susceptibility to infection and disease [296, 297].

Pulmonary TB (PTB), the most common form of the disease in South Africa, is usually diagnosed on the basis of a number of clinical features, such as persistent cough, loss of appetite, night sweats, weight loss, and chest radiography. These clinical features however, develop long after the individual has been infected by the causative organism *Mycobacterium tuberculosis* (*Mtb*), and it is believed that by the time that these clinical features become manifest, the infected individual has potentially promoted the disease epidemic by facilitating the spread of the disease to casual contacts. Furthermore, these clinical features, whilst assisting the clinician in making the final disease diagnosis, are not definitive enough, and hence clinicians also rely on the identification of *Mtb* by positive culture via the mycobacteria growth indicator tube (MGIT) system (Beckton-Dickinson, USA), and Ziehl-Neelsen (ZN) staining, and subsequent examination by light-microscopy. Thus a substantial amount of time is lost waiting for the results from the

laboratory. Therefore, the early detection of TB disease is a major priority in counteracting the disease.

4.2 The Study Cohort

Two hundred and twenty-three unrelated individuals from the SM study, ranging in age from 18 to 65 years, with a male to female ratio of 1.4:1 (58% versus 42%), were screened for *NAT2* genotypes in order to assess their acetylation status. The average age for males was 37 years and for females 33 years. All patients enrolled in the SM study gave their written informed consent of participation, and only individuals experiencing their first PTB-disease episode were enrolled. These patients, who were regularly monitored, had no indication of any other chronic disease, or pregnancy, and were HIV negative. Additional patient data, including smoking and drinking habits, was also collected via a questionnaire. All enrollees received directly observed treatment, short course (DOTs), for a period of six months, with first line therapy consisting of isoniazid, rifampicin, pyrazinamide and ethambutol, as advocated by the World health Organization (WHO).

The infecting *Mtb* strain from each patient was confirmed by ZN staining and subsequent culture (MGIT), and was also subjected to drug sensitivity analysis for susceptibility to isoniazid and rifampicin. The strain identity was determined by DNA fingerprinting, using the mycobacterial marker IS6110 [298, 299].

At enrolment (Day 0), peripheral blood was collected for genetic analyses, for assessing the total white blood cell count, as well as absolute neutrophils, monocytes, and lymphocyte content to establish the baseline measurements for each patient. Thereafter, patients were subjected to regular follow-up assessments at Day1 and 2, then weekly for two months (Day 55), and thereafter monthly, for a further period of four months. After completion of their DOTs treatment, the patients were followed-up for a further 24 months for possible subsequent episodes of TB. However, for the purposes of our acetylation study, we restricted our analysis to data collected during DOTs treatment

focussing on factors that could be predictors of risk for disease and/or possibly diminish treatment efficacy. As indicated above, the DOTs treatment period consists of 2 months of intensive treatment with a combination of four drugs (the initiation or intensive phase), followed by 4 months of treatment with a reduced drug regimen (the continuation phase). Isoniazid is an integral part of the drug regimen during both of these treatment phases.

Efficacy of treatment was assessed by monitoring patient samples for viable *Mtb* organisms via ZN-microscopy and culture to establish a "conversion time" in days, being the period from Day 0 (diagnosis) to a point midway between the last positive *Mtb* culture and two consecutive negative *Mtb* cultures [300]. In keeping with the recommendations of the International Union against Tuberculosis and Lung Disease (IUTLD) [301], we selected Day 55 as the criterion to establish treatment efficacy on the basis of whether the patient had responded to treatment ("responders") or not ("non-responders") [300, 302]

The control cohort, consisting of the 108 unrelated individuals from the same community (see chapter 3, section 3.2), had a male to female ratio of 1:1.4 (43% versus 57%). The average age for both males and females was 42 years, with an age range from 16 to 96 years. The annual rate of tuberculosis infection in this community has been reported to be greater than 3.5% [262], and whilst a recent study found a positive tuberculin skin test (TST) in 76% of those surveyed, these control individuals exhibited no overt signs of TB-disease.

All statistical analyses were performed in conjunction with the Department for Statistical Consultation, (Faculty of Health Sciences, Stellenbosch University), using the Statistical Analysis Software (SAS) version 9.1 (www.sas.com) and Stastistica, version 10 (www.statsoft.com).

4.3 Results and Discussion

4.3.1. *NAT2* SNPs in TB patients: case-control analysis

The NAT2 acetylation status was characterised by the analysis of seven single nucleotide polymorphisms (SNPs) in the *NAT2* gene, as described in a previous chapter. These SNPs occur most frequently in various populations, and have subsequently been shown to be the most informative for determining acetylation genotypes [71, 83, 168, 169].

The genotype and allele frequency distribution of these SNPs in the unaffected ("controls") and TB-affected cohorts ("cases") are indicated in Tables 4.1 and 4.2. The Hardy-Weinberg equilibrium (HWE) analysis found six of the seven SNPS analysed to be in equilibrium. Only the 803A>G SNP was not in HWE, as shown previously (discussed in Chapter 3). Furthermore, an analysis of the allele and genotype frequencies for each of the seven SNPs shows no statistically significant difference between the two cohorts (*P*-value greater than 0.05). It is interesting to note however, that the genotypes associated with the 341T>C SNP exhibit the lowest *P*-value, exhibiting the best indication of a possible trend towards an association with TB-disease. The overall similarity in the frequency distributions for these genotypes does not however support an overt association between these particular *NAT2* polymorphisms and the presence and / or incidence of TB disease per se, in the local SAC population group.

This observation is not unexpected, and supports a hypothesis that the inherent acetylation status of the individual, as defined by these SNPs, does not play a role in TB disease susceptibility or infectivity. However, the NAT enzymes are known to metabolise a number of compounds used in anti-TB treatment regimens, such as for example, PAS, metabolised by NAT1, and INH, metabolised by NAT2 [28, 44, 46]. Therefore it is likely that the impact of these enzymes will be evident in the individual's response to treatment, which may manifest as either a variation in response to treatment, or the manifestation of adverse drug reactions (ADR).

Table 4.1 Frequency distribution for the *NAT2* genotypes between the cohorts

NAT2 SNP		Genotype Frequency						Cases vs Controls
		Cases (N = 223)			Controls (N = 108)			
	0 > 1	00‡	01†	11‡	00‡	01†	11‡	*P*-value§
191	**G > A**	0.933	0.067	0	0.907	0.093	0	0.4136
282	**C > T**	0.484	0.466	0.049	0.574	0.407	0.019	0.1768
341	**T > C**	0.422	0.507	0.072	0.509	0.380	0.111	0.078
481	**C > T**	0.516	0.435	0.049	0.593	0.343	0.065	0.2657
590	**G > A**	0.655	0.314	0.031	0.741	0.250	0.009	0.1943
803	**A > G**	0.197	0.574	0.229	0.157	0.611	0.232	0.6696
857	**G > A**	0.946	0.054	0	0.917	0.083	0	0.3016

"0" and "1" represent the major and minor alleles, respectively.
‡ - indicates homozygous SNP-genotypes
† - indicates heterozygous SNP-genotypes.
§ - determined via Chi-squared (χ^2) analysis.

Table 4.2 Frequency distribution for the *NAT2* alleles analysed.

NAT2 SNP		Allele Frequency		Cases versus Controls	
		Cases, n = 446	Controls, n = 216	OR (95% CI)$	*P*-value$
191	**G‡**	431 (*97*)	216 (*95*)	1.395 (0.6159 – 3.159)	0.5143
	A	15 (*3*)	10 (*5*)		
282	**C‡**	320 (*72*)	168 (*78*)	0.7256 (0.4955 – 1.063)	0.1096
	T	126 (*28*)	48 (*22*)		
341	**T‡**	301 (*68*)	151 (*70*)	0.8936 (0.6285 – 1.271)	0.5932
	C	145 (*32*)	65 (*30*)		
481	**C‡**	326 (*73*)	165 (*76*)	0.8494 (0.5821 – 1.239)	0.4480
	T	120 (*27*)	51 (*24*)		
590	**G‡**	362 (*81*)	187 (*87*)	0.6683 (0.4229 – 1.056)	0.0982
	A	84 (*19*)	29 (*13*)		
803	**A‡**	216 (*48*)	100 (*46*)	1.089 (0.7866 – 1.509)	0.6094
	G	230 (*52*)	116 (*54*)		
857	**G‡**	434 (*97*)	207 (*97*)	1.572 (0.6521 – 3.792)	0.3463
	A	12 (*3*)	9 (*3*)		

‡ - indicates the Major Allele for each SNP
$ - Fischer's exact analysis

4.3.2. *NAT2* genotypes in TB-patients

Individual *NAT2* genotypes were scored as the summation of the results obtained for the 7 SNPs analysed, as described in a preceding chapter. Briefly, this entails the assignment of a slow acetylation (SA) status to samples possessing a combination of any two of the activity-altering SNPs (namely, 191A, 341C, 590A, and 857A), whilst the rapid acetylation (FA) status was assigned to samples without these respective polymorphisms. In addition; rapid acetylators were also scored on the basis of the presence of the 282C>T and 803A>G SNPs. The intermediate acetylation (IA) status was defined by the presence of only one of the activity-altering SNPs (Table 4.3). This convention was applied throughout this study to avoid confusion, and possible discrepancies in the assignment of the alleles.

Table 4.3 *NAT2* genotypes observed between unaffected and TB-affected individuals.

Observed Diplotype	Deduced Genotype	Unaffected Cohort		TB Cohort		P-value[§]
		N	(%)	N	(%)	
Rapid Acetylators (FF)		**24**	**(*22.2*)**	**42**	**(*18.8*)**	
0000010	*2*4/2*12A*	10	(*41.7*)	19	(*45.2*)	0.5165
0000020	*2*12A/2*12A*	6	(*25.0*)	5	(11.9)	
0000000	*2*4/2*4*	3	(*12.5*)	5	(*11.9*)	
0100010	*2*4/2*12B*	3	(*12.5*)	9	(*21.4*)	
0001010	*2*4/2*12C*	1	(*4.2*)	n.f.		
0100000	*2*4/2*13*	1	(*4.2*)	1	(*2.4*)	
0200000	*2*13/2*13*	n.f.		2	(*4.8*)	
0100020	*2*12A/2*12B*	n.f.		1	(*2.4*)	
Intermediate Acetylators (FS)		**55**	**(50.9)**	**106**	**(47.5)**	
0100110	*2*4/2*6C*	13	(*23.6*)	15	(*14.2*)	0.1074
0011010	*2*4/2*5B*	11	(*20.0*)	22	(*20.8*)	
0011020	*2*12A/2*5B*	5	(*9.1*)	24	(*22.6*)	
0010010	*2*4/2*5C*	4	(*7.3*)	5	(*4.7*)	
1100010	*2*12A/2*14B*	3	(*5.5*)	1	(0.9)	
0111010	*2*4/2*5G*	2	(*3.6*)	5	(4.7)	
0100100	*2*4/2*6A*	2	(*3.6*)	17	(16.0)	
0100001	*2*4/2*7B*	2	(3.6)	1	(0.9)	
1000010	*2*4/2*14E*	2	(3.6)	1	(*0.9*)	
0010020	*2*12A/2*5C*	2	(3.6)	5	(*4.7*)	
0100011	*2*12A/2*7B*	2	(3.6)	1	(*0.9*)	
0110010	*2*13/2*5C*	2	(3.6)	3	(*2.8*)	
1000000	*2*4/2*14A*	1	(*1.8*)	1	(*0.9*)	
1100000	*2*4/2*14B*	1	(*1.8*)	1	(*0.9*)	
0000110	*2*12A/2*6B*	1	(*1.8*)	n.f.		
0110000	*2*13/2*5D*	1	(*1.8*)	n.f		
0200100	*2*13/2*6A*	1	(*1.8*)	n.f		

0011000	*2*4/2*5A*	n.f	1 (*0.9*)	
0000100	*2*4/2*6B*	n.f	1 (*0.9*)	
0111020	*2*12A/2*5G*	n.f	1 (*0.9*)	
0110020	*2*12B/2*5C*	n.f	1 (*0.9*)	
Slow Acetylators (SS)		**29 (26.9)**	**75 (33.6)**	
0022020	*2*5B/2*5B*	7 (*24.1*)	10 (*13.3*)	
0111110	*2*5B/2*6A*	6 (*20.7*)	28 (*37.3*)	
0021020	*2*5B/2*5C*	5 (*17.2*)	2 (*2.7*)	
0111011	*2*5B/2*7B*	2 (*6.9*)	6 (*8.0*)	
1011010	*2*5B/2*14A*	2 (*6.9*)	2 (*2.7*)	
0011110	*2*5B/2*6B*	1 (*3.5*)	1 (*1.3*)	
0100200	*2*6A/2*6B*	1 (*3.5*)	2 (*2.7*)	
0200101	*2*6A/2*7B*	1 (*3.5*)	2 (*2.7*)	
0111100	*2*5A/2*6A*	1 (*3.5*)	n.f.	
0111101	*2*5A/2*7B*	1 (*3.5*)	n.f.	
0010011	*2*5C/2*7A*	1 (*3.5*)	n.f.	
1100100	*2*6B/2*14B*	1 (*3.5*)	n.f.	0.1911
0200200	*2*6A/2*6A*	n.f.	5 (*6.7*)	
1111010	*2*5B/2*14B*	n.f.	4 (*5.3*)	
0110110	*2*5C/2*6A*	n.f.	3 (*4.0*)	
0020020	*2*5C/2*5C*	n.f.	2 (*2.7*)	
0022010	*2*5A/2*5B*	n.f.	1 (*1.3*)	
0021010	*2*5B/2*5D*	n.f.	1 (*1.3*)	
1110010	*2*5C/2*14B*	n.f.	1 (*1.3*)	
1110000	*2*5D/2*14B*	n.f.	1 (*1.3*)	
0100101	*2*6A/2*7A*	n.f.	1 (*1.3*)	
1200100	*2*6A/2*14B*	n.f.	1 (*1.3*)	
1200001	*2*7B/2*14B*	n.f.	1 (*1.3*)	

Table 4.3 continued…..

§ - global *P*-value determined by the Chi-squared (χ^2) analysis.

n.f. = not found.

The frequency distribution for the FA, IA and SA genotypes is comparable between the cohorts, and the small differences between them does not correlate significantly with either the controls or cases (see Table 4.3). It is likely however, that these differences failed to achieve a significant correlation due to the relatively small sample set employed in this experiment. However, a number of interesting features are evident between the cohorts, such as a slightly greater frequency with respect to FA and IA groups in the control cohort, whilst the SA status prevails amongst the cases.

In the FA group, the *NAT2*4/2*12A* and NAT2**4/2*4* genotypes are equally distributed between the controls and the cases. Secondly, the *NAT2*4/2*12A* genotype was the most common genotype in both cohorts. In contrast to this, the *NAT2*12A/2*12A* and the *NAT2*4/2*12B* genotypes represent a greater proportion amongst the control and TB-affected individuals respectively (an approximately 2x difference between the cohorts). This is interesting, since the polymorphisms classifying these *NAT2*12* haplotypes, namely, 803G (for *NAT2*12A),* and 282T plus 803G (for *NAT2*12B*), do not confer a reduced acetylation activity [177, 303]. In effect, the 803 polymorphism changes the amino acid lysine (K) to arginine (R) at position 268 (K268R), whilst the 282C>T polymorphism is a synonymous variant, maintaining the amino acid identity at position 94, which is tyrosine (Y). In this respect, there may be some significance in the observation of an approximately two-fold difference in the prevalence of these particular *NAT2*12* alleles between the two cohorts.

Given the well-known redundancy in the genetic code, synonymous polymorphisms are generally regarded as being equivalent to one another, since they do not alter the amino acid identity in the protein. However, it has been shown that the use of the alternative codon in these so-called "silent" polymorphisms is not a random event in biological systems [304-306]. A definite correlation has for example been reported between the extent of bias in codon usage, and the level of gene expression [304, 307, 308]. This type of codon bias has also been shown in different tissues [309, 310]. More recently, 'silent" polymorphisms have also been found to influence the activity of proteins even at the level substrate specificity [265].

The prevalence of the *NAT2*4/2*12B* rapid acetylator genotype in the TB cohort may therefore indicate a potential reduction in the treatment efficacy in these individuals, possibly due to an increased rate of INH metabolism, thereby causing individuals with this genotype to require a relatively longer period of time to be cured, due to the relatively rapid removal of the drug (INH). Consequently, these individuals may be at risk of developing multi-drug resistance in the infecting *Mtb* strain, due to the impairment of the efficacy of INH in the standard treatment regimen. Whilst it may be advisable to administer an increased dosage of INH to individuals with this genotype to achieve the required treatment efficacy, the specific acetylation effect of this genotype on the

metabolism of INH, still needs to be determined empirically, since the relative merits of each of these codons in influencing the level of *NAT2* gene expression, and hence protein activity, is unknown.

In the case of the IA group, there are several genotypes that occur more frequently amongst the control samples than amongst TB-affected individuals. These differences generally range from approximately 2-fold (for *NAT2*4/2*6C* (1.7x), to as much as 6-fold (for *NAT2*12A/2*14B*). Furthermore, the *NAT2*12A/2*14B* and *NAT2*4/2*14B* genotypes respectively, have a 6-fold and 2-fold greater occurrence between the cohorts. Hence it would appear that the presence of the 803G variant (in *NAT2*12A*) may afford a 3-fold greater acetylation advantage in combination with the *NAT2*14B* allele, than would the wild type *NAT2*4* allele. However, in the case of the analogous *NAT2*12A/2*7B* and *NAT2*4/2*7B* genotypes, the difference in the frequency distribution of these particular genotypes remain unchanged between the cohorts (at 4-fold greater in the controls). This latter observation may indicate that the apparent advantageous effect of the 803G polymorphism is counteracted by different SA alleles. Alternatively, it is likely that the acetylation effect conferred by the *NAT2*14B* allele may be more pronounced, since it has been shown that the 857A SNP variant, present in the *NAT2*7* alleles, exhibits a decreased binding affinity for the cofactor acetylCoA [176, 177]. This interaction is an overriding step in the acetylation reaction, and its retardation or inhibition would preclude the subsequent acetylation of the substrate (INH). Hence the overall acetylation rate for the *NAT2*7* allele may be slower relative to that of the *NAT2*14* alleles.

Of equal interest within the IA group, particularly in the context of factors that may reduce anti-TB treatment, is the observation that the *NAT2*12A/2*5B* and *NAT2*4/2*6A* genotypes occur more frequently amongst the cases than the controls, at 2.5-fold and 4.4-fold, respectively. These genotypes may therefore have an effect on the efficacy of anti-TB treatment, and predispose individuals to a relatively longer treatment episode. Furthermore, of the 14 IA genotypes commonly observed among the cohorts, approximately 43% (6/14) occurred at a frequency of less than 1% in the TB-affected cohort. However, the Chi-squared analysis for the IA group (*P*-value = 0.1074) indicates

that there is no significant differences in the genotypes between the two cohorts, whilst some of the genotypes, such as for example, *NAT2*4/2*12A*, and *NAT2*4/2*5B* are equally distributed between them.

The most prevalent acetylation genotypes in the SA group are *NAT2*5B/2*5B* (24.1% in controls versus 13.3% in cases), and *NAT2*5B/2*6A* (20.7% in controls versus 37.3% in cases). This 2-fold difference between the cohorts with respect to these genotypes may only have a slight impact upon the difference in treatment efficacy, in those individuals harbouring the *NAT2*5B/2*6A* genotype. On the other hand, the *NAT2*5B/2*14A* and *NAT2*5B/2*5C* genotypes, being less prevalent among the cases, at approximately 3-fold and 6-fold, respectively, could be indicative of a better prognosis with respect to the treatment efficacy, and for this reason may be more effective in improving the overall disease epidemiology, by improving the rate of cure of the disease. In this respect, the trend for the SA group is similar to that observed for the FA and IA acetylation groups, in that several of the commonly occurring *NAT2* genotypes are found more frequently amongst the individuals unaffected by TB disease (controls).

Equally interesting in the SA group, is the difference in number of unique *NAT2* genotypes that occurred between the cohorts, with 28% (21/75) observed amongst the cases and only 14% (4/29) in the controls. In the case of the FA and IA groups, the number of unique genotypes are relatively small, and not that different between the cohorts, being 4.2% and 4.8% (controls versus cases) in the FA group, and 5.5% and 5.3% (controls versus cases) in the IA group. The importance of these differences among the SA group should be investigated further, on the basis of a dual genotype-phenotype assay to determine the significance of these unique genotypes with respect to anti-TB treatment efficacy.

The chi-squared analyses indicate that there are no significant differences in the FA, IA and SA states between the cohorts, and no obvious trend can be observed in the different cohort groups, beyond the general differences apparent in the frequency distribution of some genotypes between the cohorts. Nevertheless the differences observed for these

particular genotypes in the FA, IA and SA states, may warrant further investigation in those individuals to assess the role of NAT2 acetylation underlying TB disease.

A possible confounding factor may reside in the alternative allele assignments that can be deduced from observed diplotypes (Table 4.4). These assignments could not be resolved unambiguously by the experimental methodologies generally used in the *NAT2* genotyping assays [170, 173, 311]. However, we found that in most situations of such ambiguity, the alternative allele and genotype assignment(s) did not alter the overall acetylation status of the individual sample. Furthermore, this 7-SNP assay for determining *NAT2* genotypes has also been used successfully by several other authors [272, 312].

Table 4.4 Alternative NAT2 genotypes deduced from observed diplotypes

Observed Diplotype	Deduced Genotype	Proposed Alternative Genotype	Acetylation Change
Fast Acetylation			
0001010	*2*4/2*12C*	*2*11/2*12A*	-
Intermediate Acetylation			
0100110	*2*4/2*6C*	*2*12A/B/2*6A/B*	-
0011010	*2*4/2*5B*	*2*12A/2*5A*	
0010010	*2*4/2*5C*	*2*12A/2*5D*	-
1100010	*2*12A/2*14B*	*2*12B/2*14A*	-
0111010	*2*4/2*5G*	*2*13/2*5B or 2*12B/2*5C*	-
0100100	*2*4/2*6A*	*2*13/2*6B*	
0100001	*2*4/2*7B*	*2*13/2*7A*	-
1000010	*2*4/2*14E*	*2*12A/2*14A*	-
0100011	*2*12A/2*7B*	*2*12B/2*7A*	-
0110010	*2*13/2*5C*	*2*12B/2*5D*	-
1100000	*2*4/2*14B*	*2*13/2*14A*	-
0000110	*2*12A/2*6B*	*2*4/2*6F*	-
Slow Acetylation			
0111110	*2*5B/2*6A*	*2*5G/2*6B or 2*5C/2*6N*	-
0111011	*2*5B/2*7B*	*2*5G/2*7A or 2*7C/2*5A*	-
1011010	*2*5B/2*14A*	*2*4/2*14C*	Intermediate (IA)
0011110	*2*5B/2*6B*	*2*5C/2*6E or 2*5A/2*6F*	
1100100	*2*6B/2*14B*	*2*4/2*14D*	Intermediate (IA)
0111100	*2*5A/2*6A*	*2*5D/2*6N*	-
1111010	*2*5B/2*14B*	*2*13/2*14C*	Intermediate (IA)
0110110	*2*5C/2*6A*	*2*12B/2*5E*	Intermediate (IA)
0021010	*2*5B/2*5D*	*2*5A/2*5C*	-
1110010	*2*5C/2*14B*	*2*13/2*14F*	Intermediate (IA)
0100101	*2*6A/2*7A*	*2*4/2*6J*	Intermediate (IA)

Some of these alternative assignments, particularly among SA individuals, did alter the proposed genotype to IA. In all of these latter instances the change in genotype is due to the assignment of more than one of the activity-altering SNPs to a particular haplotype. For example, the observed diplotype 1011010, can be resolved as either the *NAT2*5B/2*14*A genotype (SA), or the alternative, *NAT2*4/2*14C* genotype (IA) (table 4.4). In the latter genotype, all four of the observed polymorphic SNPs (191A, 341C, 481T, and 803G) are assigned to one allele, namely *NAT2*14C*. Similar haplotype combinations (to that for *NAT2*14C*) have been reported by other authors [313-315], but the effect of these particular alleles (haplotypes) on the overall acetylation status of the enzyme has, to our knowledge, not yet been elucidated. The current practise in the acetylation research field assigns the activity-altering polymorphisms to separate alleles (see table 4.5), a protocol that we have also followed, since the experimental techniques are unable to rapidly determine the phasing of these SNPs in particular genotypes.

Comparisons using pair-wise linkage disequilibrium (LD) analyses (Haploview version 4.2, http://www.broad.mit.edu/mpg/haploview) [316] to align these 7 SNPs were unable to indicate an over-riding pattern of associations between particular SNPs. Some authors have attempted to resolve this question by using a computer programme (PHASE version 2.1.1, http://www.stat.washington.edu/stephens/software.html) for the reconstruction of haplotypes from population data [317]. These authors found that while these haplotype prediction techniques cannot fully determine the actual *NAT2* genotype, the ambiguity in prediction only varies between 0.1% to 1.2% [266, 270]. Furthermore, Agúndez (2003) [271] reported that the 282 polymorphism has not been found on the same *NAT2* allele (haplotype) as the 341 polymorphism, thereby providing some support for the argument against the generalised assignment of SNP variants to one haplotype.

Thus whilst ambiguities in *NAT2* genotyping may prevail in studies using the current methodologies, the degree of discordance may be small and may not be significantly relevant across populations due to the variation in the *NAT2* SNP frequencies between populations. The inclusion of additional, known activity-altering SNPs, such as for example 364G>A, 411A>T, 434A>C, and 845A>C, may become a major factor in reducing this ambiguity or discordance.

Table 4.5 Nucleotide alterations defining various *NAT2* alleles.

NAT Allele	Nucleotide Change							Amino Acid Change Defining Acetylation Phenotype	Phenotype
	191	282	341	481	590	803	857		
***2*4*‡**	**G**	**C**	**T**	**C**	**G**	**A**	**G**	**None**	Rapid
2*5A			*C*	*T*					Slow
2*5B			*C*	*T*		*G*			Slow
2*5C			*C*			*G*			Slow
2*5D			*C*					**I^{114}→T**	Slow
2*5E			*C*		*A*				Slow
2*5G		*T*	*C*	*T*		*G*			Slow
2*6A		*T*			*A*				Slow
2*6B					*A*			**R^{197}→Q**	Slow
2*6C		*T*			*A*	*G*			Slow
2*6J		*T*			*A*		*A*		Slow
2*7A							*A*	**G^{286}→E**	Slow
2*7B		*T*					*A*		Slow
2*12A						*G*		**K^{268}→R**	Rapid
2*12B		*T*				*G*			Rapid
2*12C				*T*		*G*			
2*13		*T*						**Y^{94}→Y** (None)	Rapid
2*14A	*A*							**R^{64}→Q**	Slow
2*14B	*A*	*T*							Slow
2*14C	*A*		*C*	*T*		*G*			Slow
2*14D	*A*	*T*			*A*				Slow
2*14E	*A*					*G*			Slow
2*14F	*A*		*C*			*G*			Slow

‡ - indicates the wild type (reference) allele.
Blocked text indicates the defining polymorphisms for the various allele classes.
Amino acid Abbreviations:- Glutamate (E); Glycine (G); Isoleucine (I); Lysine (K); Glutamine (Q); Arginine (R); Threonine (T); Tyrosine (Y).

4.3.3 *NAT2* alleles and haplotypes in TB-patients

An analysis of the NAT2 acetylation distribution with respect to the rapid and slow alleles reflects the general trend observed in the genotype analyses, in that no significant correlations were found in the overall allele frequencies between the cohorts (Table 4.6). This trend was also reflected in the allele subclasses, such as for example, the *NAT2*12* rapid alleles, and the *NAT2*5* slow alleles.

Table 4.6 Distribution of the *NAT2* alleles within the cohorts

Deduced Genotype	Unaffected Cohort		TB Cohort		*P*-Value†
	N	(%)	N	(%)	
Fast Alleles	**103**	**(47.7)**	**190**	**(42.6)**	**0.9306**
NAT2*4	**57**	**(57.3)**	**109**	**(57.4)**	
*2*12A*	37	(34.0)	62	(32.6)	0.5672
*2*12B*	3	(2.9)	11	(5.8)	
*2*12C*	1	(1.0)	n.f.		
NAT2*12	**41**	**(37.9)**	**73**	**(38.4)**	
NAT2*13	**5**	**(4.9)**	**8**	**(4.2)**	
Slow Alleles	**113**	**(52.3)**	**256**	**(57.4)**	**0.2801**
2*5A	2	(1.8)	3	(1.2)	0.4661
*2*5B*	46	(40.7)	114	(44.5)	
*2*5C*	14	(12.4)	23	(9.0)	
*2*5D*	1	(0.9)	2	(0.8)	
*2*5G*	2	(1.8)	3	(1.2)	
NAT2*5	**65**	**(57.5)**	**145**	**(56.6)**	
*2*6A*	12	(10.6)	64	(25.0)	0.0026*
*2*6B*	4	(3.5)	5	(2.0)	
*2*6C*	13	(11.5)	15	(5.9)	
NAT2*6	**29**	**(25.6)**	**84**	**(32.9)**	
*2*7A*	1	(0.9)	1	(0.4)	1.000
*2*7B*	8	(7.1)	11	(4.3)	
NAT2*7	**9**	**(8.0)**	**12**	**(4.7)**	
*2*14A*	5	(4.4)	4	(2.0)	0.4422
*2*14B*	5	(4.4)	10	(3.9)	
*2*14E*	n.f		1	(0.4)	
NAT2*14	**10**	**(8.8)**	**15**	**(6.3)**	

† indicates the global *P*-value determined by Chi-squared analysis or Fischer's exact analysis.
* - *P*-value < 0.05 represents a statistically significant association.

However, a significant association was observed for the *NAT2*6* class of alleles (*P*-value = 0.0026) which indicates that these particular slow alleles may have a potential impact on TB epidemiology. We also tested this association between particular *NAT2*6* alleles, namely *NAT2*6A* versus *NAT2*6C*, since these two haplotypes only differ on the basis of the 803G SNP, present in *NAT2*6C*. In the latter instance, the significance of the correlation was confirmed (*P*-value = 0.0035). In contrast, the frequency of the *NAT2*6B* allele did not differ substantially between the cohorts. Therefore, the significant correlation observed for the *NAT2*6* alleles between controls and cases is associated with differences in the *NAT2*6A* and *NAT2*6C* alleles.

As indicated in table 4.5, *NAT2*6* alleles share the 590A polymorphism, which imparts a slow acetylation activity to the enzyme. However, the greater prevalence of particular haplotypes in controls compared to cases (*NAT2*6C*) and vice versa (*NAT2*6A)* would appear to indicate that these additional *NAT2* SNPs may afford differing acetylation capacities to the individuals in these groups. Hence the *NAT2*6A* allele (haplotype) may have a greater confounding effect on TB treatment efficacy, than *NAT2*6C*, due to the synonymous SNP at 282. Alternatively, the rapid acetylation status conferred by the 803G SNP variant may impart a beneficial affect with respect to the overall acetylation activity.

In this context it may be interesting that the *NAT2*6C* allele occurred among IA individuals and not among SA individuals in the SAC cohorts. Among the IA individuals the *NAT2*4/2*6C* genotype was associated with the control cohort more than the TB-affected cohort. On the other hand, the *NAT2*4/2*6A* genotype showed a greater association with TB-affected IA individuals. Furthermore, in the SA group, the *NAT2*6A/2*6A* genotype was not found among control individuals, while the *NAT2*6A/2*6B* genotype occurred equally among control and TB-affected SA individuals.

4.4 The Surrogate Marker surveillance study

The surrogate marker surveillance study has the added advantage of providing specific demographic data pertaining to TB disease, since the patients were strictly monitored during their DOTs treatment episode. In this respect this study has the potential

to broaden the understanding of the impact of the acetylation status on the epidemiology of this very complex disease. Table 4.7 and 4.8 list some of the demographic data collected to promote an understanding of the ramifications implicit to the disease. It is envisaged that this information may facilitate the attainment of certain goals which may assist in the curtailment of the disease epidemic.

Table 4.7 Demographic data for the SM cohort[§]

Demographic		Count (n = 223)	Frequency (%)
Gender	Male	130	58.3
	Female	93	41.7
Age (years)	<20	8	3.6
	21-30	83	37.2
	31-40	63	28.3
	41-50	44	19.7
	>51	25	11.21
INH[S]	Yes	23	10.3
	No	197	88.3
	Unknown	3	1.4
Education	None	5	2.2
	Primary level	160	71.8
	Secondary level	48	21.5
	Tertiary level	2	<1
	Unknown	8	3.6
Monthly Income	None	16	7.2
	<R250	22	9.9
	R251-1000	28	12.6
	R1001-2000	22	9.9
	> R2000	6	2.3
	Unknown	129	57.9
Consume Alcohol	Yes	171	76.7
	No	36	16.1
	Unknown	16	7.2
Smoke	Yes	199	89.2
	No	12	5.4
	Unknown	12	5.4
BMI[†]	Male average	18.34[†]	
	Female average	18.32[†]	
Height[*]	Male average	1.654	
	Female average	1.653	

INH[S] = Indicates *Mtb* strains sensitive to isoniazid.
[†] = body mass index.
[*] = Height measured in metres.
[§] = adapted from Babb, C.L. (2007) Ph.D. Dissertation.

Table 4.8 *Mtb* strain families identified the SM study cohort

Strain Family[Δ]	Distribution frequency N	(%)	Median Sputum Smear Conversion Day	Median Culture Conversion Day
0	6	2.7	21.5	67.5
2	8	3.6	42	60
4	4	1.8	72.5	65.5
6	6	2.7	21	42.8
7	7	3.1	23	58.5
8	1	0.4	21	41.5
9	5	2.2	41	56
11	**47**	**20.9**	**42**	**72**
12	1	0.4	139	42.5
13	3	1.3	21	42
14	6	2.7	23	31.5
16	4	1.8	31	66.5
17	1	0.4	10	73.5
18	4	1.8	27.5	57.8
19	1	0.4	73	59
21	2	0.9	23	47.2
23	2	0.9	41	21
24	2	0.9	42.5	42
25	1	0.4	2	20.5
27	2	0.9	57.5	57.3
28	6	2.7	20.5	41.3
29	**69**	**30.7**	**41**	**72**
30	1	0.4	212	116
31	1	0.4	72	72
110	5	2.2	22.5	43.5
120	8	3.6	42	56.8
130	2	0.9	57	72
140	**18**	**8.0**	**38**	**43**
150	2	0.9	16	42.3

[Δ] = *Mtb* strain family nomenclature according to Richardson *et al* [318]

Bold type indicates the most frequently occurring *Mtb* strains.

4.4.1 Assessing strain prevalence in the cohort

As indicated in table 4.8, DNA fingerprinting analysis identified 29 different *Mtb* strains among the 223 SM cases, of which 45% (13/29) occurred at a frequency less than 1%. The most prevalent *Mtb* strains were identified as belonging to family 29 (Beijing clade), family 11 (LAM clade) and family 140 (Low copy-number clade), which co-incidentally are also the most prevalent infecting strains in the Western Cape Province of South Africa [298, 319]. This may be an indication of "strain success" for family 29 and 11 which together cause 51.6% of the disease burden in this community, whereas family 140, the next most prevalent *Mtb* strain, only accounts for 8.0% of the disease prevalence in this cohort. Alternatively, the most dominant *Mtb* strains may have become "established" in this community due to their relatively longer association with it.

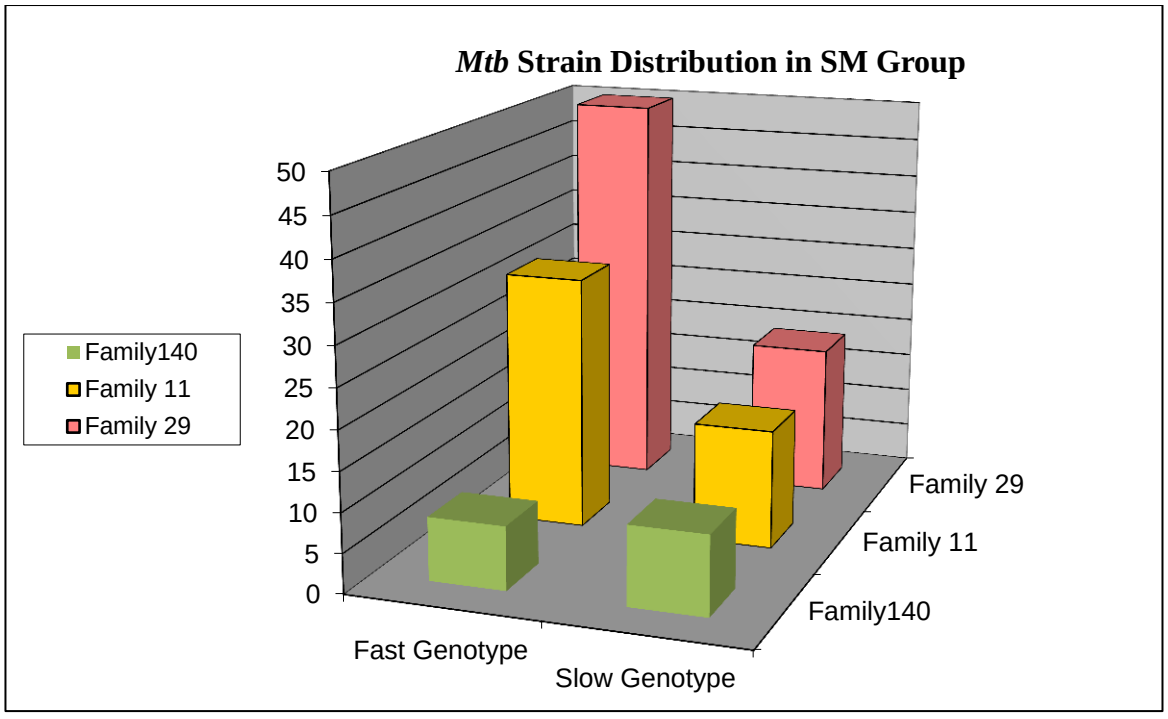


Figure 4.1 Sixty percent of the identified *Mtb* strains belong to only three strain families. The x-axis represents the *NAT2* acetylation genotypes, whilst the y-axis represents the number of individuals harbouring a specific strain family. No significant association was observed between the strains and the rapid (FA, IA) or slow acetylation (SA) activity groups. (*P* value = 0.0777)

A chi-squared analysis to investigate a possible correlation between the presence of the most prevalent infecting strains and acetylation status as described by the rapid (FA and IA) versus slow (SA) acetylation found an association of borderline significance only (*P*-value = 0.0777) (Fig. 4.1).

It is probable that this association may be due to the fact that these strains are over-represented in this community, yet this in itself is significant, since *Mtb* strains appear to have a niche host-population relationship [6, 7, 320]. The significance of this relationship is also evident in a report which found that a particular *Mtb* strain can be an independent risk factor for treatment failure in some populations [321]. Moreover the Beijing strain has been found to associate with TB-susceptible host genetic polymorphisms in the SCL11A1/NRAMP (natural resistance-associated macrophage protein 1) gene in the local SAC population [322], as well as in an Indonesian population [323]. Therefore, it would be interesting to investigate the associative trend observed in this study, between rapid acetylation and the *Mtb* strain families 11 and 29, in a future investigation, given that family 29 also belongs to the Beijing clade.

4.4.2 Assessing treatment outcomes via sputum conversion

The SM patients were also investigated for various host-derived factors that could predict treatment outcome based on the comparison of ZN-microscopy (smear) and culture analysis of patient sputa at enrolment and after 2 months of DOTs treatment, according to the National Treatment Programme of South Africa. Early data from the SM cohort indicated a difference in the sputum results for some patients after 8 weeks of treatment, in that the majority of these patients (91) recorded a negative sputum result, and were labelled as the responder group (R), whilst 21 patients recorded a positive sputum result after 8 weeks of treatment. These latter patients were referred to as the delayed or Non-responder (NR) group.

A cohort of 36 patients (18 responders and 18 non-responders), matched for age, gender, extent of disease (number of cavities as well as the extent of disease-affected alveolar), sputum smear grade (bacterial load in sputum sample) and good adherence to treatment (completed 32 of the required 40 treatment doses (DOTs)) were subsequently selected

and investigated for their NAT2 acetylation status as well as for differences in immunological biomarkers. The results of the latter analysis was reported by Brahmbhatt *et al* [293], who observed differences in several immune markers, such as soluble tumour necrosis factor alpha receptors 1 and 2 (sTNF-R1and sTNF-R2), soluble interleukin-2 receptor alpha (sIL-2Rα), and the absolute counts of selected white blood cells (such as monocytes and neutrophils) which could be used to predict whether a patient would exhibit a delayed response to treatment (be a NR) or not.

Only 34 patients of this matched SM group were assessed for their acetylation, since the DNA from one sample of the responder group could not be amplified. This sample pair was therefore removed from the analysis. Table 4.9 summarises the acetylation results observed for this group of patients.

Table 4.9 *NAT2* acetylation polymorphisms for the matched group of SM patients

Observed Diplotype	Deduced Genotype	Sputum ZN-smear Conversion	
		Responders N (%)	Non-responder N (%)
Rapid Acetylators (FF)		**6 (*35.3*)**	**2 (*11.8*)**
0200000	*2*13/2*13*	2 (*11.8*)	n.f.
0000010	*2*4/2*12A*	1 (*5.9*)	1 (5.9)
0100010	*2*4/2*12B*	1 (*5.9*)	1 *(5.9)*
0100000	*2*4/2*13*	1 (*5.9*)	n.f.
0000020	*2*12A/2*12A*	1 (*5.9*)	n.f.
Intermediate Acetylators (FS)		**2 (*11.8*)**	**14 (*82.4*)**
0011020	*2*12A/2*5B*	n.f.	6 (*35.3*)
0100100	*2*4/2*6A*	n.f.	3 (*17.6*)
0011020	*2*4/2*6C*	n.f.	2 (*11.8*)
0011010	*2*4/2*5B*	1 (*5.9*)	1 (*5.9*)
0010010	*2*4/2*5C*	1 (*5.9*)	n.f.
1000000	*2*4/2*14A*	n.f	1 (5.9)
0010020	*2*12A/2*5C*	n.f.	1 (5.9)
Slow Acetylators (SS)		**9 (*52.9*)**	**1 (*5.9*)**
0111110	*2*5B/2*6A*	3 (*17.6*	n.f
0200200	*2*6A/2*6A*	2 (*11.8*)	n.f.
0111011	*2*5B/2*7B*	2 (*11.8*)	n.f.
0111011	*2*5C/2*5C*	1 (*5.9*)	n.f.
0100200	*2*6A/2*6B*	1 (*5.9*)	n.f.
0020020	*2*5B/2*5C*	n.f.	1 (*5.9*)

n.f. = not found

A highly significant association was observed between the acetylation status (FA, IA and SA) and the observed difference in the sputum ZN-smear conversion response between the responders (R) and Non-responders (NR) (*P*-value = 0.0002, (chi-square analysis)). The majority of the responders (88%) belonged to the SA and the FA classes respectively. These results indicate that SA and FA responded to the DOTs treatment, with a better treatment efficacy evident amongst the slow (53%) than rapid (35%) acetylators. On the other hand the observed delay in treatment efficacy, as measured by sputum response after 8 weeks of treatment, could be attributed to the effect of the IA status in the NR patients.

Given these initial results for this matched cohort, we extended our acetylation analysis to include all patients from the SM study cohort who had completed the initiation phase of treatment, and for whom both ZN-smear and culture results were available. The median conversion time for ZN-smear was 41 days (range 1-231days), and for culture 71.5 days (range 0.5-268.5days) respectively. At Day 55, the data for 213 patients was audited for treatment efficacy only, and the patients were classified as either Responders or Non-responders, as delineated by ZN-smear and culture conversion results. For this analysis we defined the Non-responders as those individuals who exhibited a positive ZN-smear and culture result at Day 55. Of these patients 213 patients, 190 were sensitive to isoniazid (INH^S), and 21 were resistant (INH^R), whilst 2 were of unknown drug sensitivity. However all of the patients with INH^R *Mtb* strains returned negative culture and ZN-smear results at Day 55 (hence classified as responders), were included in the analysis. The overall acetylation results for this cohort of samples is summarised in Table 4.10.

Table 4.10 NAT2 acetylation in Responder and Non-responder patients.

Acetylation	ZN-Smear Conversion[§]				Culture Conversion[§]			
	Responder		Non-Responder		Responder		Non-responder	
Rapid (FA)	30	*19.2*	11	*19.3*	20	*22.2*	21	*17.1*
Intermediate (IA)	77	*49.4*	26	*45.6*	43	*47.8*	60	*48.8*
Slow (SA)	49	*31.4*	20	*35.1*	27	*30.0*	42	*34.1*
Total	**156**	***100%***	**57**	***100%***	**90**	***100%***	**123**	***100%***

[§] = no significant association found for either ZN-Smear (*P*-Value = 0.863) nor Culture conversion (*P*-Value = 0.606).

As indicated in the table 4.10 the acetylation status does not differ between the cohorts. However, upon assessing the prevalence of the individual *NAT2* SNPs in the responder versus the non-responder groups several differences become evident (Tables 4.11 and 4.12).

Table 4.11 *NAT2* genotypes in the responder and non-responder patient groups

NAT2 SNPs		ZN-Smear Conversion (n = 213)			Culture Conversion (n = 213)		
		R n = 156 (%)	NR n = 57 (%)	*P*-value[‡]	R n = 90 (%)	NR n = 123 (%)	*P*-value[‡]
191	**GG**	149 (*95.5*)	49 (*86*)	**0.0298**	86 (*95.6*)	120 (*97.6*)	0.4592
	GA	7 (*4.5*)	8 (*14*)		4 (*4.4*)	3 (*2.4*)	
282	**CC**	81 (*51.9*)	24 (*42.1*)	0.2771	51 (*56.7*)	87 (*70.7*)	0.1023
	CT	66 (*42.3*)	31 (*54.4*)		34 (*37.8*)	32 (*26.0*)	
	TT	9 (*5.8*)	2 (*3.5*)		5 (*5.6*)	4 (*3.3*)	
341	**TT**	65 (*41.7*)	26 (*45.6*)	0.5314	39 (*43.3*)	83 (*67.5*)	**0.0015**
	TC	79 (*50.6*)	29 (*50.9*)		43 (*47.8*)	36 (*29.3*)	
	CC	12 (*7.7*)	2 (*3.5*)		8 (8.9)	4 (*3.3*)	
481	**CC**	81 (*51.9*)	31 (*54.4*)	0.4712	48 (*53.3*)	90 (*73.2*)	**0.0089**
	CT	66 (*42.3*)	25 (*43.9*)		36 (*40.0*)	30 (*24.4*)	
	TT	9 (*5.8*)	1 (*1.8*)		6 (*6.7*)	3 (*2.4*)	

	GG	102 (*65.4*)	38 (*66.7*)		62 (*68.9*)	97 (*78.9*)	
590	**GA**	48 (*30.8*)	18 (*31.6*)	0.7501	26 (*28.9*)	22 (*17.9*)	0.1588
	AA	6 (*3.8*)	1 (*1.7*)		2 (*2.2*)	4 (*3.3*)	
	AA	30 (*19.2*)	11 (*19.3*)		18 (*20.0*)	12 (*9.8*)	
803	**AG**	90 (*57.7*)	32 (*56.1*)	0.9782	52 (*57.8*)	95 (*77.2*)	**0.0094**
	GG	36 (*23.1*)	14 (*24.6*)		20 (*22.2*)	16 (*13.0*)	
	GG	151 (*96.8*)	52 (*91.2*)		86 (*95.6*)	122 (*99.2*)	
857				0.1362			0.1648
	GA	5 (*3.2*)	5 (*8.8*)		4 (*4.4*)	1 (*0.8*)	

Table 4.11 continued…

R = Responder; NR = Non-responder.
‡ = Statistical results obtained by Fischer's exact or Chi-Squared (χ^2) analyses (Results in bold indicate statistically significant observations).

Table 4.12 *NAT2* SNP genotypes associated with Culture conversion

***NAT2* SNP**		**Genotype Frequency**			
		Responders	**Non-responders**	**OR (95% CI)**[$]	***P*-value**[$]
341	**TT**	39	83	0.3934 (0.2194 – 0.7054)	**0.0020**
	TC	43	36		
	TC	43	36	0.5972 (0.1661 – 2.147)	0.5394
	CC	8	4		
	TT	39	83	0.2349 (0.0667 – 0.8278)	**0.0247**
	CC	8	4		
481	**CC**	48	90	0.444 (0.244 – 0.808)	**0.0096**
	CT	36	30		
	CT	36	30	0.6000 (0.1382 – 2.606)	0.7228
	TT	6	3		
	CC	48	90	0.2667 (0.0638 – 1.114)	0.0752
	TT	6	3		
803	**AA**	18	12	2.740 (1.225 – 6.130)	**0.0145**
	AG	52	95		
	AG	52	95	0.43790 (0.2090 – 0.9173)	**0.0357**
	GG	20	16		
	AA	18	12	1.200 (0.4490 – 3.207)	0.8047
	GG	20	16		

[$] = Fischer's exact analysis. **Bold** type indicates significant association (*P*-value less than 0.05).

In the case of the ZN-smear conversion (table 4.10), the heterozygous GA genotype at SNP 191 is significantly associated with non-responders (NR) rather than responders thereby indicating that this genotype may have a confounding effect on the treatment efficacy in the NR group. None of the genotypes of the other 6 SNPs investigated show any correlation with either of these groups in the ZN-smear conversion cohort.

In the culture conversion assessment, a significant correlation was observed for the *NAT2* genotypes delineated at positions 341, 481 and 803 respectively. In the case of the 341- and 481-SNPs the wild-type genotypes (341TT and 481CC, respectively) are significantly associated with NR, whereas in the case of the 803-SNP, the heterozygous variant (803AG) is significantly associated with the NR group.

Approximately 58% (123/213) of patients were classified as Non-responders by their culture conversion result at Day 55, despite the observation that the majority of the *Mtb* strains (89.4%) in this group were sensitive to isoniazid. This would indicate that the response to treatment involves a greater complexity with several other factors, such as the interplay between the host's immune response and the virulence of the mycobacterium, playing an equally important role [324]. Another factor concerns the viability of the mycobacteria in lung lesions, which ranges from actively growing, - hence conducive to the antibacterial activity of the drugs - to fully dormant and intractable to drugs [325, 326]. Since INH is known to exert its anti-mycobacterial activity against actively growing organisms, the viability of the bacteria in the lung lesions would play an important role in the actual survival of the organism in the presence of INH.

4.5 Conclusion

The effect of *NAT2* gene polymorphisms was investigated as a possible confounding factor to anti-TB treatment efficacy in South African Coloured (SAC) individuals from a high TB-incidence community. A seven single nucleotide polymorphism (SNP) assay was used to delineated the enzyme activity on the basis of rapid (FA), intermediate (IA), and slow (SA) acetylation, and in the initial case-control association analysis no acetylation differences were observed between cases and controls with respect to the above activity classifications. However, a number of interesting trends were observed with respect to particular *NAT2* genotypes and alleles, which may merit further investigation.

In the FA class the *NAT2*4/2*12B* genotypes appear to associate more with the cases rather than controls, whilst in the IA class, a similar trend is observed for the *NAT2*12A/2*5B* and *NAT2*4/2*6A* genotypes. It is interesting that one of the common features associated with these above genotypes involves synonymous SNPs (282C>T and 481C>T), which are generally not considered to affect the acetylation activity of the enzyme. Various publications have however, reported that synonymous variants can exert significant effects on protein activity by influencing the expression level of genes [265, 304-306], possibly via ribosome stalling [327]. This would result in aberrantly folded proteins which, are removed via the proteasome-ubiquitin pathway [112], thereby depleting the pool of active protein. This aspect of synonymous polymorphisms has not as yet been investigated with respect to the various known haplotype combinations present in *NAT2*. Whilst the results observed for the FA, IA, and SA groups were not significant, possibly due to the relatively small cohort sample, some of the trends observed among particular allele subclasses, may become significant in a larger study cohort. This may constitute an interesting avenue for future research.

In the case of the SA alleles, the *NAT2*6A* allele was significantly associated with TB-affected individuals. This allele results in an impaired acetylation activity, which may be further compounded by the 282C>T synonymous polymorphism, and may be a confounding factor with respect to anti-TB treatment efficacy. A recent publication by Higuchi *et al* (2007) [243] found a significance association between the *NAT2*6A*

haplotype and the development of drug-induced hepatotoxicity (DIH) in Japanese individuals. In this respect, DIH would have an important negative effect on TB-treatment efficacy, since its appearance in TB patients is idiosyncratic, and necessitates the suspension of treatment [328, 329]. Given the genetic ancestry of the local SAC community, particularly with respect to their oriental antecedents [245], the *NAT2*6A* allele (haplotype) may be an important biomarker in TB-affected individuals.

Our experimental analysis also investigated the acetylation genotype or status in a well defined group of TB patients from the SM cohort. These individuals were strictly monitored and assessed for environmental and host-derived demographic factors that could affect the treatment efficacy during the intensive phase of treatment (Day 55, month 2), as recommended by the International Union against Tuberculosis and Lung Disease (IUTLD) [301]. On the basis of their treatment results at this time point, these patients were classified as either responders, who gave a negative result, or non-responders, who scored positive for *Mtb*, as assessed by both ZN-smear and culture analyses.

Our results indicated that FA, IA and SA activity did not differ between responders and non-responders, which imply that the acetylation activity per se does not influence the treatment efficacy during the intensive phase of treatment (Day 55). However, a number of significant correlations were observed for certain *NAT2* SNPs and the observed treatment response in the ZN-smear conversion and culture conversion groups, respectively. The 191GA *NAT2* genotype was significantly associated with the non-responders in the Zn-smear conversion group, yet did not show a similar correlation in the culture conversion assessment. Furthermore, none of the other *NAT2* SNPs showed any correlation in the ZN-smear conversion group. On the other hand, in the culture conversion group, only the genotypes delineated by the 341,- 481,- and 803-SNPs were significantly associated with non-responders. These results indicate that only certain particular *NAT2* genotypes may have a confounding effect on treatment efficacy as assessed by ZN-smear or culture conversion.

Individually, these SNPs define the separate *NAT2*14* (191), *NAT2*5* (341), *NAT2*11* (481) and *NAT2*12* (803) alleles or haplotypes, which exhibit disparate effects on the acetylation activity of the enzyme as explained in an earlier chapter. In this context, these

individual effects may underlie the singularity of the observations in the ZN-smear and culture conversion assessments. Alternatively, these SNPs collectively define the *NAT2*14C* allele, or the *NAT2*5/2*14* SA genotype. Given the technology currently available for genotypic analysis, it is not yet possible to delineate the effects of the separate SNPs as observed for the ZN-smear and culture conversion assessments.

The majority of the patients assessed by culture conversion were non-responders, who exhibited a delay in converting to culture negative, indicating that there are factors other than acetylation activity affecting the response to treatment. Interestingly, we were able to show a slight borderline association between the infecting *Mtb* strain (family 29 and family 11) and the rapid acetylation genotype. However this associative trend warrants further investigation to establish its significance.

In an analogous genetic study of the SM cohort, Babb *et al* [300] reported several factors that could be correlated with treatment efficacy. These authors found that the extent of disease (as measured by degree of the lung affected by TB disease and number of cavities) and the total number and type of white blood cells at diagnosis were predictive of the ZN-smear and culture conversion times in the responder / non-responder patients. Furthermore smoking and SNPs in the Vitamin D receptor also predicted treatment response, as measured by ZN-smear conversion time. These factors did not show a similar association with culture conversion time. Similarly, Brahmbhatt *et al* [293] indicated that the differences in certain serum biomarkers and white blood cell counts, as measured at diagnosis (Day 0) and after 4 weeks of treatment could predict sputum ZN-smear status after 8 weeks of treatment. Carroll *et al* [330] and Hesseling *et al* [302] also found that the delay in ZN-smear conversion was linked to a higher bacterial load in the sputum sample at diagnosis, which in turn may be indicative of more extensive TB-disease, or greater virulence of the infecting mycobacterial strain. These results indicate that the extent of disease may be a major indicator of treatment response during the initiation phase of DOTs treatment.

These reports indicate that several host derived factors, such as the immunological response to the mycobacterial infection, as well as mycobacterial “virulence” may contribute to the extent of the disease manifestation. Secondly, this may be exacerbated

by individual habits (such as smoking, and delays in seeking medical treatment) and social stigma, which may contribute to the disease becoming established in the host. Given that the DOTs treatment regimen is standardised on the basis of milligram drug per kilogram body weight of the patient [301], coupled to inherent, yet largely neglected acetylation activity if the individual's NAT2 enzyme, the above factors may collectively facilitate the observed delay in the curative treatment response. However, given the proven effectiveness of INH against active, drug sensitive *Mtb* strains (as found in the SM study cohort), a better treatment efficacy could be facilitated by adjustments to the standardised DOTs treatment regimen, given a prior knowledge of the individual's acetylation status.

CHAPTER 5

An investigation of *NAT1* Polymorphisms in South African populations

5.1 Introduction

The action of drug metabolising enzymes in response to environmental insults is important, particularly given their involvement in the biotransformation of therapeutic drugs and xenobiotics. From a health perspective, there is a mounting interest in understanding the action of these enzymes due to the increased burden of iatrogenic illnesses stemming from therapeutic drug usage [331, 332]. In this context, the study of arylamine *N*-acetyltransferases (NAT1 and NAT2) set a historic precedent. These conjugating enzymes are responsible for the *N*-acetylation, *O*-acetylation, and *N*, *O*-acetylation of various amine-containing aromatic compounds [68, 80, 182], via a complex series of reactions, in metabolic pathways that result in the detoxification or activation of these substrates (Fig. 5.1)

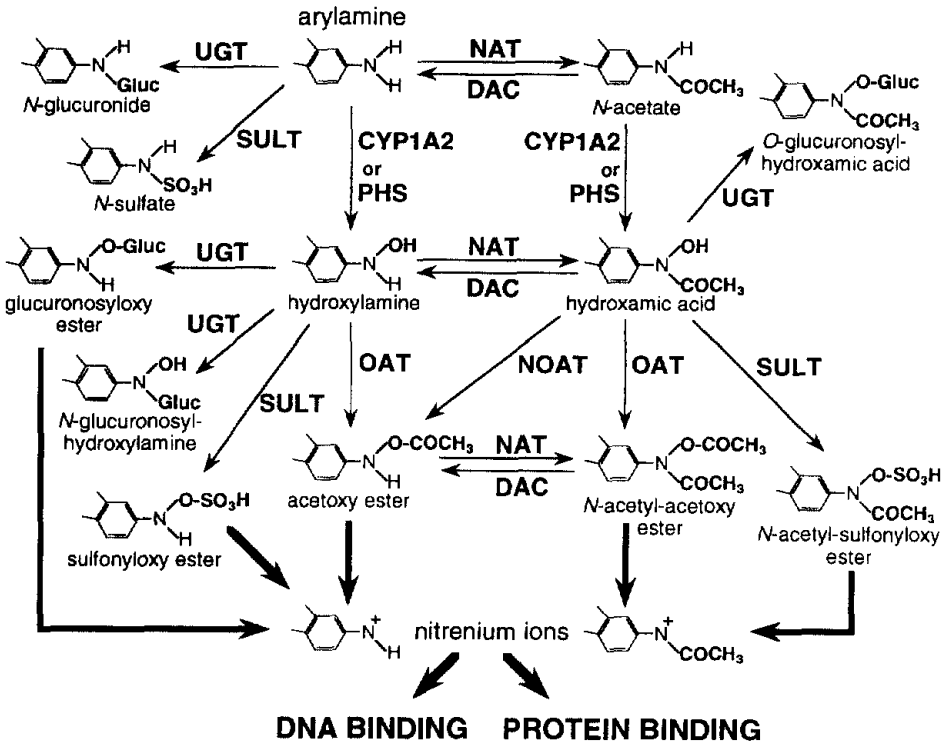


Figure 5.1. Proposed pathways for the metabolic biotransformation of amine-containing aromatic substrates, from Grant *et al*, 1997 [182]. The enzyme abbreviations are: UGT, UDP-glucuronosyltransferase; NAT, arylamine N-acetyltransferase; DAC, microsomal deacetylase; SULT, sulphotransferase; CYP1A2, cytochrome P4501A2; PHS, prostaglandin H synthase; OAT, *O*-acetyltransferase; NOAT, intramolecular *N*,*O*-acetyltransferase (OAT and NOAT reactions are catalysed by the NATs).

The NAT enzymes play an important part in these metabolic pathways, therefore are potentially linked to sources of toxicological susceptibility [333]. Whilst it is well-known that these enzymes have separate substrate specificities [28, 66-68], there is tantalising evidence to suggest that there is an inter-dependence between these enzymes during acetylation. Mark *et al* [206] has shown that *para*-aminobenzoic acid (pABA), a known NAT1 substrate, is formed as a by-product of the metabolism of procainamide (PRO), a known NAT2-substrate, thereby suggesting that these enzymes may act sequentially. Furthermore, studies by Du Souich *et al* [45], and Dutcher *et al* [334], have indicated that PRO and acetyl-PRO, a subsequent by-product, are metabolised independently of each other, whilst Cribb *et al* [335] has implicated both NAT2 and NAT1 in the metabolism of caffeine. More recently, studies by Kawamura *et al* [97] have shown that certain substrates that are metabolised via acetylation, such as, 4-ethoxyaniline (EOA), 5-aminosalicyclic acid (5AS), and 4-iodoaniline (IOA) are equally acetylated by both NAT1 and NAT2.

A recent study high-lighted the importance of detailing the genetic characterization of African populations in order to adapt treatment regimens for these patients [167]. This becomes all the more significant given the high burden of diseases, such as tuberculosis (TB) (World Health Organization (WHO) Factsheet No. 104, November 2010, http://www.who.int/mediacentre/factsheets/fs104/en/), and the distinctive ethnic and genetic diversities on the African continent, where extensive genetic variation exists even amongst populations located geographically close to one another [336]. More recently, micro-array based comparative genome hybridization studies indicate that population differences in expressed genetic phenotypes can be correlated with population clustering [337, 338]. Hence it is important for countries, particularly developing countries, to research the genetic challenges underpinning local disease epidemiology. It is within this context that we investigated *NAT1* genetic polymorphisms within local populations, given that 4-aminosalicyclic acid (4AS or PAS), a well-known component of treatment regimens of multi-drug resistant (MDR) *Mtb*, is metabolised by NAT1 [44].

5.2 The Sample Sources

DNA samples from the unrelated healthy, Caucasian and SAC individuals described in chapter 3 (section 3.2) were also investigated for polymorphisms in the 19 exonic and 3 intronic *NAT1* SNPs, as detailed previously (Table 1.1, Chapter 1). Secondly, we also tested the samples from the Surrogate Marker study cohort (chapter 4 section 4.2) for these *NAT1* SNPs, because of the overlapping substrate specificity and inter-dependence in metabolism between the NAT isoenzymes.

5.3 Results and Discussion

The *NAT1* gene sequence was initially amplified by a PCR assay using primers 1 and 2 (Table 5.1) to yield the 1492 bp sequence domain of interest. This amplified domain was then sequenced at the Central Analytical Facility, Stellenbosch University, using primers 2, 3 and 4.

Table 5.1 Primers used in the analysis of the *NAT1* gene

Primer	DNA Sequence	Tm*
1. NAT1/117	5'-^{117}AGGATACCAGTTGGAATCTCTCT	54.8°C
2. NAT1/1609	5'-^{1609}AACAATAAACCAACATTAAAAGCTT	50.5°C
3. NAT1/409	5-^{409}GCTTTCGTTTTGTTTTCC	47.6°C
4. NAT1/899	5'-^{899}GTGTCTTCCGTTTGAC	47.4°C

Tm* = Annealing temperature, as determined by supplier (http://www.idtdna.com)

Table 5.2 summarises the 22 *NAT1* single nucleotide polymorphisms (SNPs) investigated in the sample cohorts. We obtained sequencing results for 349 samples (of 359 sent for analysis), consisting of 50 Caucasian, 109 SAC, and 190 SM individuals.

Table 5.2 List of the *NAT1* nucleotides analysed for genetic polymorphisms.

NAT1 SNP	Nucleotide Position in gene‡	Rs Identifier	Amino acid Change	Phenotype
***21T>G*†**	**461**	**4986992**	**L7L**	**Wild-Type§**
97C>T	**537**	**56318881**	**R33Stop**	**None***
190C>T	**630**	**56379106**	**R64W**	**Slow**
350G>C	**790**	**72554606**	**R117T**	**Unknown**
351G>C	**791**		**R117T**	**Unknown**
***402T>C*†**	**842**	**-**	**P134P**	**Wild-Type§**
445G>A	**885**	**4987076**	**V149I**	**Rapid**
***459G>A*†**	**899**	**4986990**	**T153T**	**Wild-Type§**
497G>C	**937**	**72554608**	**R166T**	**Unknown**
498G>C	**938**		**R166T**	**Unknown**
499G>C	**939**		**R167Q**	**Unknown**
559C>T	**999**	**5030839**	**R187Stop**	**Slow**
560G>A	**1000**	**4986782**	**R187Q**	**Slow**
613A>G	**1053**	**72554609**	**M205V**	**Wild-Type§**
640T>G	**1080**	**4986783**	**S214A**	**Unknown**
752A>T	**1192**	**56172717**	**D251V**	**Slow**
***777T>C*†**	**1217**	**4986991**	**S259S**	**Wild-Type§**
781G>A	**1221**	**72554610**	**E261K**	**Wild-Type§**
787A>G	**1227**	**72554611**	**I263V**	**Wild-Type§**
884A>G	**1324**	**55793712**	**None**	**Unknown**
1088T>A	**1528**	**1057126**	**None**	**Unknown**
1095C>A	**1535**	**15561**	**None**	**Unknown**

Information for the table was derived from the NAT consensus website. (http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/)
§ = refers to the reference *NAT1*4* allele.
† = SNPs coding for a synonymous amino acid change.
* = truncated protein, no enzyme activity.
‡ = nucleotides numbered in accordance with the reference sequence (Genbank Accession Number AJ307007).
Amino acid Abbreviations define - Aspartate (D), Glutamate (E), Isoleucine (I), Lysine (K), Leucine (L), Methionine (M), Proline (P), Glutamine (Q), Arginine (R), Serine (S), Threonine (T), Valine (V) and Tryptophan (W).

The Caucasian group had a median age of 74 years (range 51 to 86) and consisted of 18 males and 32 females, whilst the SAC group had a median age of 30 years (range 17 to 59) of whom 22 were male and 80 female. In the SM group the median age was 33 years (range 18 to 64) with 112 males and 78 females.

The results of the sequencing reactions were compared to the *NAT1* reference sequence AJ307007 (Genbank Accession Number), and DNA sequence analysis was performed using the computer programme Sequencher, version 4.10.1 (Gene Codes Corporation, USA (http://www.genecodes.com)).

5.3.1 *NAT1* polymorphisms in Caucasians

No polymorphisms were observed in 15 of the 22 exonic SNP positions analysed in this group. Furthermore, only the homozygous wild-type (WT) genotype (allele) was observed at these SNP positions. Single nucleotide polymorphisms were however observed for the exonic positions at 190C>T, 445G>A, 459G>A, 560G>A and 640T>G, respectively, where only the heterozygous genotype (allele) was observed (Table 5.3).

Table 5.3 *NAT1* polymorphisms in Caucasian individuals

Sample No.	Age	Sex	190C>T	445G>A	459G>A	560G>A	640T>G	1088T>A	1095C>A	*NAT1* Genotype
C1	75	F	C/T					T/T	C/C	*1*4/1*17*
C2	80	F	C/T					T/T	C/C	*1*4/1*17*
C3	70	M		G/A	G/A		T/G	T/T	C/C	*1*4/1*11*
C4	77	M		G/A	G/A		T/G	T/T	A/A	*1*3/1*11*
C5	57	M		G/A	G/A		T/G	A/A	A/A	*n.d.*
C6	79	F				G/A		T/A	C/A	*1*4/1*14A*
C7	85	F				G/A		A/A	A/A	*1*10/1*14A*
C8	72	F				G/A		T/A	C/A	*1*10/1*14B*
C9	78	F						T/A	A/A	*1*3/1*10*
C10	76	F						T/T	A/A	*1*3/1*3*

n.d.= “not defined”, indicating a SNP combination that could not be given a *NAT1* allele classification.

Interestingly, these results were obtained in only 8 individuals of the group (16%), 5 of whom only possessed a single polymorphism in the *NAT1* exonic region. These SNPs confer the slow acetylation activity on the enzyme, and classify the *NAT1*17* (190C>T) and *NAT1*14* (560G>A) alleles respectively [74, 110, 339]. The *NAT1*11* alleles, classified by the 445G>A SNP, is observed in two of the remaining 3 samples, whilst in the remaining sample the observed SNP combination could not be classified as a *NAT1* allele according to the NAT consensus database. The effect of the *NAT1*11* alleles on the acetylation activity of this enzyme has yet to be characterised [340, 341]. In addition to the above 8 samples, a further 2 samples showed intronic SNPs only. Whilst both of these samples have the homozygous 1095AA polymorphism, which represents the *NAT1*3/1*3* genotype, the other sample, classified as *NAT1*3/1*10*, has the additional 1088CA heterozygous genotype. These 10 samples account for 20% of the *NAT1* genotypes observed in the Caucasian group.

The highest frequency distributions were observed for the 1088T>A and 1095C>A intronic SNPs, with their wild-type alleles, 1088T and 1095C respectively, being the most prevalent in the group (71% and 70% respectively). These SNPs together define the *NAT1*10* allele, which was previously considered to be a gain of function allele, exhibiting a more enhanced acetylation capacity than the reference *NAT1*4* allele [148, 158]. However, according to the NAT consensus website, the phenotypic activity of these alleles is currently listed as "unknown". Nevertheless, in addition to the reference allele (*NAT1*4*), these SNPs account for 80% of the *NAT1* alleles observed in this group (discussed later).

5.3.2 *NAT1* polymorphisms in SAC individuals

In the case of the SAC group, a similar pattern was observed to that described above for the Caucasian group, in that the greatest frequency distribution for polymorphic variants was again observed for the 1088T>A and 1095C>A SNPs. However, in this group the 1088T and 1095C alleles occurred at a relative frequency of only 0.53.

Table 5.4 summarises the exonic SNP variants observed in the SAC group. Despite showing polymorphic variants at 7 SNP positions, two more than was observed in the

Caucasian group, only one of the SNPs, namely 560G>A, (which defines the *NAT1*14* allele), has been characterised as a loss of function allele, and exhibits an acetylation activity lower than the reference allele *NAT1*4* [341]. Of the 6 remaining SNPs, the 459G>A (*NAT1*11*) and 777T>C (*NAT1*23* and *NAT1*27*) SNPs are synonymous variants [342], whilst the acetylation activity conferred the other 4 SNPs has yet to be determined empirically (NAT consensus website). We also observed a number of samples containing SNP combinations that could not be classified as *NAT1* alleles according to the information in the NAT website. These alleles ("not defined") would require further investigation and characterisation, particularly since they have SNPs (459G>A and 640T>G) which encode amino acid changes that could potentially alter the activity of the enzyme.

Table 5.4 *NAT1* polymorphisms in SAC individuals

Sample No.	Age	Sex	445G>A	459G>A	497G>C	498G>C	560G>A	640T>G	777T>C	1088T>A	1095C>A	*NAT1* Genotype
SAC1	30	F	A/A	A/A				T/G		T/A	C/A	*n.d.*
SAC2	28	M		G/A						T/T	C/C	*1*4/1*11*
SAC3	46	F			G/C					A/A	C/A	*n.d.*
SAC4	39	F				C/C	A/A			T/T	C/C	*n.d.*
SAC5	41	F							T/C	T/A	C/A	*1*10/1*23*
SAC6	41	M							T/C	T/T	C/A	*1*4/1*23*
SAC7	28	F								T/A	C/C	*n.d.*
SAC8	32	F								A/A	C/A	*n.d.*
SAC9	31	F								A/A	C/A	*n.d.*
SAC10	27	F								T/A	A/A	*1*3/1*10*
SAC11	22	F								T/A	A/A	*1*3/1*10*
SAC12	32	M								T/T	C/T	*1*4/1*3*
SAC13	29	F								T/T	C/T	*1*4/1*3*
SAC14	40	M								T/T	C/T	*1*4/1*3*

n.d. = "not defined", indicating SNP combinations that could not be given a *NAT1* allele classification.

5.3.3 *NAT1* polymorphisms in SM individuals

NAT1 analysis of samples from the SM group, exhibited polymorphic variants at 13 of the 22 SNPs investigated, 10 exonic and 3 intronic SNPs (Table 5.5). Five *NAT1* SNPs occurred exclusively in this group, namely 21T>G (*NAT1*27*), 350G>C, 351G>C (*NAT1*5*), 781G>A (*NAT1*24*), 787A>G (*NAT1*25*), and 884A>G (*NAT1*5*). The *NAT1*24*, *1*25* and *1*27* alleles encode an acetylation activity equivalent to *NAT1*4* [339, 341], whilst the activity of the *NAT1*5* allele has yet to be characterised [70].

As in the previous groups the *NAT1*14* allele (560G>A) was the only loss of function allele observed in this group that could have an effect on the NAT1 acetylation activity [341, 343]. Furthermore, as indicated before, there are several SNPs whose effect on the acetylation activity has yet to be determined. Interestingly, some of these SNPs and alleles also occurred in the Caucasian group. The significance of this observation needs further investigation, since some of these alleles, such as for example, *NAT1*5* (35G>C, 351G>C) and *NAT1*11* (640T>G) encode amino acid changes that could have some potential significance on NAT1 acetylation activity (see Table 5.2).

It was not possible to classify several of the SNPs to particular *NAT1* alleles, mainly because these observed SNP combinations are not described in the *NAT* consensus database (http://louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/). As indicated in Table 5.5, several of these assignments involve homozygous allele variants for the SNPs at positions 1088 and 1095 in combination with one or more of the exonic SNPs. For example, in the case of samples SM12 and SM13, which possess the 781AA (*NAT1*24* allele) and 787GG (*NAT1*25* allele) homozygous SNPs respectively, these alleles have not been described in combination with SNPs at positions 1088 and 1095 (in the NAT consensus database). Therefore, these SNP combinations which we observed in this population represent previously undescribed alleles and haplotypes. These SNPs account for 12.9% (14/109) of the *NAT1* genotypes in this group, and a significant proportion of this group (6/14 (42.9%)) could not be assigned to a known *NAT1* haplotype (alleles), and would therefore constitute an interesting avenue for future research.

As in the previous groups, the most frequent polymorphisms were again observed for the intronic SNPs at NAT 1088 and 1095 respectively, with an observed frequency for the 1088T allele equal to 0.505, whilst the 1095C allele occurred at a frequency of 0.457. These frequencies were lower than observed in the Caucasian group, and may indicate a potential activity difference between these ethnic groups, since the *NAT1*10* allele (1088A plus 1095A polymorphisms) has previously been found to be a rapid acetylator using para-aminobenzoic acid as a probe drug [76, 158]. However, Hein *et al* [344] and Zhu and Hein [341] have shown that the NAT enzymes exhibit differing acetylation activities towards different substrates, and in this context, the activity of these *NAT1*10* alleles would need to be re-evaluated.

Table 5.5 *NAT1* polymorphisms in SAC individuals from the SM study group

Sample No.	Age	Sex	21T>G	350G>C	351G>C	445G>A	459G>A	560G>A	640T>G	777T>C	781G>A	787A>G	884A>G	1088T>A	1095C>A	*NAT1* Genotype
SM1	24	M	T/G											T/A	C/A	*n.d.*
SM2	36	F	T/G											T/T	C/C	*n.d.*
SM3	28	F		C/C	G/G									A/A	A/A	*n.d.*
SM4	60	M		G/C				G/A		T/C				T/T	C/C	*n.d*
SM5	53	M					G/A							T/T	C/C	*1*4/1*30*
SM6	61	M				A/A	A/A		T/G					A/A	A/A	*n.d.*
SM7	21	F				G/A	G/A		T/G					A/A	A/A	*n.d*
SM8	37	M				G/A	G/A		T/G					A/A	A/A	*n.d*
SM9	25	M				A/A	A/A		G/G		A/A			T/T	A/A	*n.d*
SM10	28	M	T/G							T/C				T/T	C/C	*1*4/1*27*
SM11	32	M								T/C				T/A	C/A	*n.d.*
SM12	33	F										G/G		A/A	A/A	*n.d*
SM13	25	F									A/A			A/A	A/A	*n.d.*
SM14	33	M								T/C				T/T	C/C	*1*4/1*23*
SM15	63	M								T/C				T/T	C/C	*1*4/1*23*
SM16	39	M								T/C				T/A	C/A	*1*10/1*23*
SM17	25	F											G/G	T/A	C/A	*n.d.*
SM18	31	M												T/A	A/A	*1*3/1*10*
SM19	38	MN												A/A	C/C	*n.d.*
SM20	42	F												A/A	C/A	*n.d.*
SM21	34	M												A/A	C/C	*n.d.*
SM22	36	M												A/A	C/A	*n.d.*
SM23	31	M										G/G	G/G	T/A	C/A	*n.d.*

SM24	33	F												T/T	C/A	*1*4/1*3*
SM25	59	M												T/T	C/A	*1*4/1*3*
SM26	26	F												T/T	C/A	*1*4/1*3*
SM27	46	F	T/G							T/C				T/A	C/A	*1*10/1*27*
SM28	24	M												T/A	A/A	*1*3/1*10*

…. *Table 5.5 continued*..

n.d. = “not defined”, indicating SNP combinations that could not be classified as known *NAT1* alleles.

Underlined nucleotides represent novel bases observed at the respective SNP position(s).

The most prevalent genotypes observed in the three study groups were *NAT1*4/1*4*, *NAT1*4/1*10*, and *NAT1*10/1*10*, and together these account for between 80 and 87% of the acetylator alleles (genotypes) described in these groups. The relative frequency distribution for these genotypes is summarised in Tables 5.6 and 5.7.

Table 5.6 Frequency distribution for the *NAT1* genotypes in the study groups

***NAT1* Genotype**	**Genotype Frequency**					
	Caucasian		**SAC**		**SM group**	
1*4/1*4	27	(*54%*)	29	(*26.6%*)	50	(*26.3%*)
1*4/1*10	8	(*16%*)	42	(*38.5%*)	65	(*34.2%*)
1*10/1*10	5	(*10%*)	24	(*22.0%*)	47	(*24.7%*)
‡***NAT1****	10	(*20%*)	14	(*12.8%*)	28	(*14.7%*)
TOTAL	**50**	**(*100%*)**	**109**	**(*100%*)**	**190**	**(*100%*)**

Table 5.7 Comparison of *NAT1* genotype frequency between Caucasian and SAC individuals

***NAT1* Genotype**	**Genotype Frequency**					
	Caucasian group		**SAC group§**		***P*-value**	**OR (CI) †**
1*4/1*4	27	(*54%*)	79	(*26.4%*)	<0.0001	3.236 (1.862-5.625)
1*4/1*10	8	(*16%*)	107	(*35.8%*)		
1*10/1*10	5	(*10%*)	71	(*23.7%*)		
‡***NAT1****	10	(*20%*)	42	(*14.1%*)		
TOTAL	**50**	**(*1000%*)**	**299**	**(*100%*)**		

† = statistical analysis via Fischer's exact test

‡ = *NAT1* genotypes that do not contain the *1*4* or *1*10* alleles

§ = combined group consisting of individuals from both the healthy and TB-affected groups (SAC and SM).

The results in table 5.6 indicate that the different *NAT1* genotypes have a comparable frequency distribution between the healthy SAC individuals and the TB-affected individuals from the SM study group. Since these individuals all belong to the South African Coloured ethnic group, these observations indicate that these *NAT1* genotypes are not correlated with TB disease in this ethnic group. On the other hand, the frequency of these genotypes were very different in the Caucasian group, where the *NAT1*4/1*4* genotype occurred at a relative frequency of 54%. The *NAT1*4* confers rapid NAT1 acetylation activity, and this allele is classified as the reference allele. The acetylation activity of *NAT1*10* on the other hand is controversial, since the SNPs defining this haplotype occur in the 3'-untranslated region of the gene, and hence it is thought that these SNPs would not have an effect on NAT1 enzyme activity.

Table 5.8 shows the allele distribution observed in this study for the Caucasian and SAC individuals in comparison to similar results reported for different ethnic group. It is evident from the table that the *NAT1* allele frequencies observed for the South African Caucasian group is comparable to that observed for Caucasians in Europe. In the case of the SAC individuals the overall pattern is quite similar to that previous reported for South African Blacks [155]. On the other hand the relative frequency for the *NAT1*10* allele is similar to that previously reported for Thai and Japanese individuals. Hence these NAT1 results lends support to the idea that the SAC community has received genetic contributions from Oriental antecedents [245].

Table 5.8 *NAT1* allele distribution as reported for different ethnic groups

Ethnic Group	No. of Alleles tested (n)	Allele frequency expressed as a percentage (%)				
		1*4	***1*3***	***1*10***	***1*11***	***Other****
Caucasian						
British[155]	224	73.7	3.6	18.7	3.1	0.9
German[76]	628	70.9	3.0	20.1	3.3	2.7
African						
Lebonese[156]	84	56	3.6	10.7	-	29.8
S.A. Black[155]	202	48.5	1.0	50.5	-	-
Asian						
Indian[154]	280	51	30	17	2	-
Thai[201]	466	50.4	3.4	43.8	2.4	-
Japanese[161]	98	52.6	1.0	40.6	-	5.7
RSA‡						
Caucasian	100	62	4	18	3	13
SAC	598	44.3	1.7	41.6	0.2	12.2

‡ = Results observed in this study of local ethnic groups.

Other* = *NAT1*14*, *NAT1*15*, *NAT1*18*, and Unassigned alleles, not yet listed on the NAT1 website.

5.5 Conclusion

We have investigated *NAT1* polymorphisms in ethnic populations in Cape Town South Africa in order to characterise the prevalent genetic variants that may influence the NAT1 acetylation activity. Twenty-two *NAT1* SNPs and alleles, described in the NAT consensus database, were analysed via automated DNA sequencing. Our results indicate SNP variants for 16% of Caucasian samples, 6% of SAC and 9% of the SM group samples analysed. Furthermore, in the majority of these samples only a single exonic locus was observed to be polymorphic.

The commonly occurring SNPs in all three groups were at positions 445, 459, 560 and 640, and of these, only the 560G>A SNP has been characterised as a loss of function

allele (slow acetylator). However, this SNP occurs rarely in the SAC individuals, being observed at a frequency of less than 1% and hence would not be considered to make an appreciable contribution to the change in the overall NAT1 acetylation activity in this population. In the Caucasian group, this SNP had a frequency of occurrence of 6%.

The highest level of polymorphic variation involved the 1088T>A and 1095C>A intronic SNPs, and these observations are similar to results recently reported by Mortensen *et al* (2011) [345] for diverse groups of African and non-African populations. The effect of these SNPs and hence their relevance to NAT1 acetylation is somewhat controversial at present. The *NAT1*10* allele is defined by the 1088T>A and 1095C>A SNP combination, and various studies have reported an association between the *NAT1*10* allele and the occurrence of various cancers [76, 147, 148]. Furthermore, this *NAT1*10* allele has been associated with rapid acetylation activity in several cancer studies, such as gastric [149], larynx [346], prostate [148], pancreas [347] and colorectal cancer [348]. However, since these polymorphisms occur in the 3'-untranslated region of the gene it is considered by popular convention that these polymorphisms would not alter the structure, and subsequently the activity function of the NAT1 protein. Therefore this feature is currently an area for future research.

In this context it is perhaps more interesting to note that no polymorphisms were observed for the 97C>T, 402T>C, 499G>C, 559C>T, 613A>G and 752A>T SNPs in any of the groups studied, which could indicate that these SNPs may possibly be omitted from the *NAT1* genotypic analysis in these ethnic groups. Two of these SNPs, namely, 402T>C (*NAT1*20*) and 613A>G (*NAT1*21*) confer rapid acetylation activity, whilst 559C>T (*NAT1*15*) and 752A>T (*NAT1*22*) are loss of function alleles, conferring slow acetylation activity. Similarly, the 97C>T SNP (*NAT1*19*) encodes a truncated NAT1 protein, which could also be considered to be a loss of function allele. Hence, given the result that several of the polymorphic alleles observed in our study are currently of unknown phenotypic status, the characterisation of these allelic variants would be an important avenue for future research since these alleles could have a substantial effect on overall NAT1 acetylation activity in our local populations.

CHAPTER 6

Rapid Detection of *NAT2* polymorphisms by the SNaPshot® technique

Rapid detection of *NAT2* polymorphisms by the SNaPshot® technique.

Marinus Barnard[#], Cedric J. Werely[#], Annemie M. Jordaan, Robin Warren, Paul D. van Helden, and Thomas C. Victor.
[#]Joint First Authors.

DST/NRF Centre of Excellence in Biomedical Tuberculosis Research / MRC Centre for Molecular and Cellular Biology, Department of Biomedical Sciences, Faculty of Health Sciences, Stellenbosch University, South Africa.

(*Unpublished Results*)

Précis

Several *NAT2* genetic polymorphisms have been validated for assessing the phenotype of the enzyme. This analysis is frequently done via a restriction enzyme - gel electrophoresis technique (RFLP analysis), which is relatively inexpensive and easy to perform, yet has the disadvantage of being laborious and time-consuming. Hence there is a need to investigate other more rapid techniques, which may be of particular relevance in a resource-limited setting, such as in most of Africa, where access to DNA sequencing or TaqMan® technology is limited or non-existent.

We therefore investigated the suitability of the SNaPshot® assay. This method enables the simultaneous analysis of multiple genetic sites, via a relatively simple yet specific primer extension assay. Using SNaPshot® we were able to discriminate both homozygous and heterozygous nucleotide alterations (alleles) at each of the genetic targets (loci), via a simple two-tube multiplex reaction assay. A significant saving in time compared to the RFLP assay was realised. Furthermore, the technique facilitated the rapid classification of the acetylation status for individual samples, in a higher throughput setup.

Our experimental design used a dual-tube multiplex assay, which obviated primer cross-reactivity during the SNaPshot® assay. We found this primer cross reactivity to be a major limitation in a single-tube assay format, using the *NAT2* gene-specific primers of relatively short lengths (less than 30 nucleotides). Increasing the length of the primers would improve primer-annealing and may solve this problem. However longer primers are more prone to forming secondary structures (hairpin loops) and/or annealing to non-specific loci, thereby generating no primer extension results or non-specific results. These factors would reduce the specificity of the assay, which is ultimately more significant than the advantage in time-saving. These limitations notwithstanding, this dual-tube SNaPshot® assay may have some advantages over the RFLP technique in a resource-limited setting, despite its limitations.

Sample Source:- Adult samples obtained from the TB clinics in the Boland-Overberg region of the Western Cape Province. Sample cohort used in this study only.

All authors contributed to the experimental design; MB and CJW contributed equally to the experimentation and analysis of results (SNaPshot® and RFLP); MB and CJW co-wrote the initial manuscript; MB and CJW are listed alphabetically.

6.1 Abstract

Background The human *NAT2* gene metabolises isoniazid, an important therapeutic drug used in anti-TB treatment regimens. Numerous *NAT2* genetic polymorphisms alter the activity of the enzyme, and can be used to classify individuals into fast, intermediate, and slow acetylators.

Design The SNaPshot® technique enables multiple single nucleotide polymorphisms to be assayed simultaneously via a primer extension assay. We used this assay to facilitates the accurate detect of the major phenotype-altering polymorphisms in the *NAT2* gene, to facilitate the characterisation of the individual's acetylation status.

Results Both homozygous and heterozygous nucleotide changes were easily discriminated at each individual locus analysed. The technique facilitated the rapid analysis of multiple targets in the gene, in a high throughput setup.

Conclusion The accurate classification of the acetylator status for individuals on anti-TB treatment regimens can reduce the incidence of adverse drug effects and promote better patient compliance, thereby improving the treatment efficacy.

Key words Primer extension, *N*-acetyltransferase, polymorphisms, pharmacogenetics, SNaPshot®

6.2 Introduction

The arylamine *N*-acetyltransferases (NAT1 and NAT2; E.C. 2.3.1.5) are important enzymes in the bio-transformational processes in the liver. These phase II drug metabolizing enzymes (DMEs), catalyse the acetyl-Coenzyme A mediated conversion of important therapeutic xenobiotics, via their reactive oxygen and nitrogen substituent atoms [68, 80, 349]. Numerous single nucleotide polymorphisms (SNPs) have been described in both the *NAT1* and *NAT2* genes, and certain SNPs within each of these genes have been associated with a reduced enzyme activity [168, 181, 342, 350]. This facilitates the speedy classification of the individual's acetylation status without the constraints of using of different probe drugs specific for NAT1 and NAT2 [68, 80]. These polymorphisms contribute to the individual's predisposition to developing certain environmentally associated diseases and various cancers [351-353]. Furthermore, acetylation studies in different population groups indicate that some of these SNPs exhibit an ethnic bias in their frequency of occurrence [63, 185], and may therefore confer alleles responsible for the inter-ethnic and inter-individual variability in acetylation.

Previous studies have shown that the *NAT2* genotyping assays accurately reflects the phenotype of individuals, and also facilitates the sub-classification of individuals into rapid, intermediate and slow acetylators [62, 170, 174]. As such, the human *NAT2* gene has been shown to be the enzyme responsible for the inactivation of INH, the primary drug used to treat TB patients. A significant number of patients on anti-TB treatment experience INH-related adverse drug reactions (ADR), which are often life-threatening [173, 213, 354]. Therefore, an *a priori* knowledge of the patient's NAT2 acetylation status could facilitate the individualisation of the INH dosages used in therapeutic anti-TB regimens, ensuring a better rate of treatment compliance, as well as reducing the incidence of drug-induced liver injury (DILI) [211].

Currently, more than 50 different alleles have been described in the *NAT2* gene (http://www.louisville.edu/medschool/pharmacology/NAT.hmt). The majority of these alleles possess combinations of only a few specific SNPs, namely, 191G>A, 282C>T, 341T>C, 481C>T, 590G>A, 803A>G, and 857G>A (Figure 4.1). Some of these SNPs encode synonymous amino acids (282C>T, and 481C>T), or conserved amino acids (803A>G), and do not reduce the activity of the enzyme. However, reduced acetylation is

conferred by the 191G>A, 341T>C, 590G>A, and 857G>A SNPs, which respectively designate the *NAT2*14*, *2*5*, *2*6*, and *2*7* allele classes. Consequently, these particular SNPs have become the ones most often investigated with respect to *NAT2* genotyping studies.

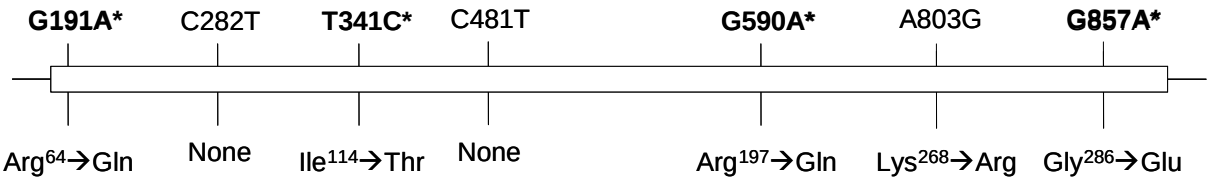


Figure 6.1 Schematic illustration of the *NAT2* gene showing the analysed SNPs. The relative SNP positions within the *NAT2* gene are indicated above the schematic, whilst their deduced amino acid changes are indicated below (http://www.louisville.edu/medschool/pharmacology/NAT2.html).
Asterisk (*) indicates alleles conferring a reduced acetylation activity.

Most of the *NAT2* genotyping assays utilise a Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) technique [355], because of the simplicity, ease, and relatively low cost of this analysis. However, this methodology is time consuming and relatively labour intensive. Subsequently a number of faster genotyping methods have been evaluated, including fluorescence resonance energy transfer (FRET) [356], WAVE DNA fragment analysis [357], and the use of Taqman Fluorogenic probes [358]. In addition, the 341T>C SNP has been genotyped using Nested PCR and Reverse Line Blot Hybridization methodologies.[359, 360]. All these methods have their inherent advantages and disadvantages, and often certain important SNPs are omitted from the genotypic analysis [171, 172]. This obviates comparisons of the SNP allele frequencies in different ethnic populations, which may be particularly significant in some populations in the context of the increasing worldwide TB epidemic, since the greatest burden of disease occurs in some geographic populations and sub strata of these groupings (http://www.who.int/mediacentre/factsheet/fs104/en/index.html).

The peculiarities of TB disease epidemiology are currently under renewed scrutiny, with greater emphasis being placed upon understanding the reasons behind issues of disease latency, re-infection, and patient non-compliance to treatment [214, 216]. In this context, it would be important to invest in a rapid *NAT2* genotyping technique, which would have the capability of analyzing the significant activity altering SNPs in a high throughput setup. This knowledge would enable the treatment regimen, particularly with respect to the INH dosage, to be adjusted for individual patients, to ensure optimal efficacy, as well as reduce the incidence of treatment interruption due to DILI. We therefore assessed the SNaPshot® technique as a means to facilitate the rapid, high throughput classification of the acetylation status of individuals, by characterising the major phenotype-altering SNPs in the *NAT2*.

6.3 Materials and methods

This study forms part of a longitudinal research project undertaken in the Western Cape Province of South Africa, amongst patients attending the TB clinics outside of the Cape Town Metropole. The project was reviewed and sanctioned by the Institutional Review Committee of Stellenbosch University (registered under the project code 2000/C061). Blood samples from randomly selected patients attending these clinics were processed for DNA via our standard laboratory protocol, as previously described [361]. All subjects who participated gave their informed consent for inclusion in the study.

PCR-Amplification A 1000 bp domain of the *NAT2* gene was initially amplified by PCR using gene specific primers previously described by Hickman *et al* [62] (Table 6.1). The 100 µl PCR reactions were carried out on a Gene Amp® System 9700 temperature cycling platform (Applied Biosystems, USA), consisting of 100 ng template DNA, 1x PCR reaction buffer (Qiagen, Germany), 3.5 mM $MgCl_2$, 4.0 mM of each of dATP, dCTP, dGTP and dTTP, 25 pmol of each primer, and 0.5 units HotStar Taq DNA polymerase I (Qiagen, Germany). The PCR amplification conditions consisted of an initial incubation at 95°C for 15 min., followed by 45 cycles of 94°C for 1 min., 53°C for 1 min., and 72°C

for 1 min., followed by a final polymerisation at 72°C for 10 min. PCR amplification products were verified in 1.5% agarose, using 1x TBE (pH 8·3) running buffer, a current of 100 v/cm for 1 hour, and visualized by ethidium bromide staining (Chapter 2).

SNaPshot® *Analysis* This technique (Applied Biosystems, USA) utilises individual primers to bind to the sequence-specific DNA domain, immediately 5'to the SNP in question. Subsequently, the primer is extended by the incorporation of a sequence-specific fluorescently-tagged ddNTP complementary base. Seven primers, varying between 17 to 22 bp in length, were designed for the primer extension (PE) analysis of the most frequently genotyped polymorphisms in the *NAT2* gene (Table 6.1).

Table 6.1 ***NAT2*** **Primers used in the genotype analysis.**

NAT2 Primers	Primer Sequence (5'-3')	Primer Length (nt)	*NAT2* Allele	Tm
PCR Primers[a]				
Nat-Hu14 (F)	5' GACATTGAAGCATATTTTGAAAG-3'	23		60ºC
Nat-Hu16 (R)	5' GATGAAAGTATTTGATGTTTAGG-3'	23		60ºC
Multiplex 1 Primers				
Msp191	5'-TCACATTGTAAGAAGAAACC-3'	20	G/A	48.9ºC
Mb341	5'-CACCTTCTCCTGCAGGTGACCA-3'	22	T/C	61.2ºC
BamH 857	5'-CTCGTGCCCAAACCTGGTGATG-3'	22	G/A	60.7ºC
Multiplex 2 Primers				
Fok282	5'-CAATGTTAGGAGGGTATTTTTA-3'	22	C/T	48ºC
Kpn481	5'-AAGAGAGAGGAATCTGGTAC-3'	20	C/T	50.5ºC
Taq590	5'-ATTTACGCTTGAACCTC-3'	17	G/A	46.6ºC
Dde803	5'-GGTTGAAGAAGTGCTGA-3'	17	A/G	49.2ºC

[a] = *NAT2* PCR primers, according to Hickman *et al* [62].
nt = nucleotide.
Tm = Melting point temperature (primer annealing).

The primers were designed using the software programme Primer Premier v5.00 (Premier Biosoft International, USA), and assessed for suitability in a multiplex reaction format, using Autodimer v1.0 [362], and Oligo Analyser software v3.0 (http://www.idtdna.com/analyzer/applications/oligoanalyzer/).

Each PE reaction contained 3 µl (0.30 ng/µl) of purified *NAT2* gene product, 5 µl of the SNaPshot ready mix, consisting of fluorescently labelled ddNTPs, Amplitaq DNA polymerase, and 1x reaction buffer, provided in the (ABI PRISM® SNaPshot® Kit, Applied Biosystems, USA), and 1 µl of the primer mix (Table 6.1), to a total volume of 10 µl. These mini-sequencing reactions were executed by 27 cycles of denaturation at 96°C for 10 sec., annealing at 50°C for 5 sec., and polymerisation at 60°C for 30 sec.

In this study two separate multiplex PE reactions were employed, consisting of the following primer combinations: primer set 1: *Taq* 590, *Fok* 282, *Kpn* 481 and *Dde* 803, and primer set 2: *Msp* 191, Mb 341 and *Bam H* 857 (Table 1). Unincorporated ddNTPs were removed by enzymatic degradation with 1unit of Shrimp Alkaline Phosphatase (SAP) (1 U/µl) at 37°C for 60 min, followed by the heat inactivation of the SAP enzyme at 75°C for 30 min. The fluorescently labelled PE products were fractionated in a single capillary system on an ABI PRISM® 3100-AVANT Genetic Analyzer, and the data was scored with the Genotyper® Software v3.7 NT, according to the manufacturers instructions. An internal size-standard (LIZ-120, Applied Biosystems) was included in every lane to ensure accurate sizing of the PE fragments.

We also analyzed these samples via our standard PCR-RFLP methodology for comparison [361]. Briefly, this entailed the specific restriction of aliquots of the *NAT2* PCR product with *Msp*I, *Fok*I, *Kpn*I, *Taq*I, *Dde*I, and *BamH*I according to the manufacturer's recommendations. These enzymes analyse the 191G>A, 282C>T, 481C>T, 590G>A, 803A>G, and 857G>A SNPs respectively. The *Msp*I, *Kpn*I and *Bam*HI restriction profiles were subjected to a further round of digestion with *Pst*I to facilitate the analysis of these restriction enzyme profiles via 5% polyacrylamide gel electrophoresis (mini PROTEAN II apparatus, BioRad, USA) with silver staining [250].

The 341T>C SNP was detected by AS-PCR, using the confronting two-pair primer method [251].

6.4 Results and Discussion

The SNaPshot® analysis clearly discriminated between homozygous and heterozygous nucleotide bases at the targeted SNP position (Fig.6.2). Ease of classification is facilitated by the incorporation of the specific complementary ddNTP base, as well as the differential migration of the SNaPshot® primer products. This is largely determined by the secondary structure of the primer, which is a property of the specific primer sequence. Hence primers of the same length are easily discriminated in the electrophoretograms. Representative electrophoretograms of the two multiplex reactions for a single patient are illustrated in Figure 6.2.

We observed complete concordance in the genotyping results obtained via the SNaPshot® and PCR-RFLP methodologies. Table 6.2 depicts the typical results observed for a few individual samples. The latter method is however more time-consuming and relatively labour intensive. Furthermore, we found that despite analysing the SNP RFLP profiles via polyacrylamide gel electrophoresis, the 191G>A SNP was sometimes prone to misclassification, since the homozygous-and heterozygous "mutant" alleles could not be assigned unambiguously. It thus became necessary to re-analyse this particular SNP via bidirectional sequencing, adding additional costs and a further delay in determining the individual's acetylation status.

We determined the *NAT2* acetylation genotype via a dual multiplex SNaPshot® reaction assay to assess the 7 informative SNPs. This was necessary to obviate primer cross-reactivity during the SNaPshot® reaction, which, could not be avoided due to the nature of the gene-specific sequences in primers of these relatively short lengths. It may be possible to overcome this problem by designing longer SNaPshot® primers greater than 30 bases in length. This would improve the target specificity of the primers by annealing to a larger region of the gene sequence, and hence may enable one to assay the different SNPs in a

single multiplex reaction. However, cross reactivity between primers, as well as the formation of secondary structures (hairpin loops), which are a property of the specific sequences, would inhibit the SNaPshot® PE reaction, and therefore should be determined empirically to assess the viability of the single multiplex reaction assay. Other studies have reported analysing as many as 14 different SNPs using the SNaPshot® multiplex reaction format [363], which confirms the viability and robustness of the technique.

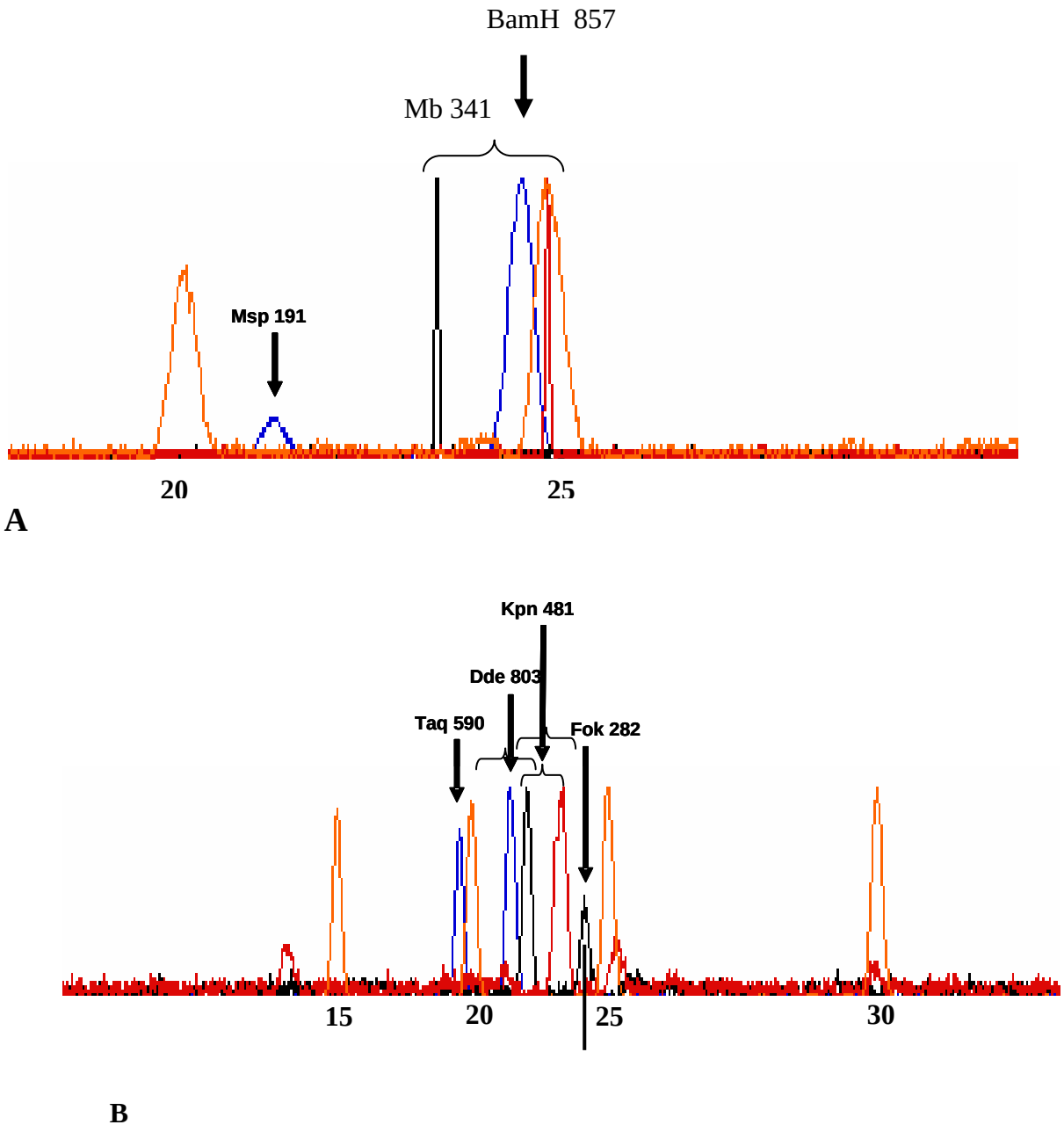


Figure 6.2 Representative electrophoretogram of the SnapShot analysis using multiplex PCR reaction mixes 1 (A) and 2 (B)
Homozygous alleles at position 191 (G/G) and 857 (G/G) are shown as single peaks; double peaks represent heterozygous alleles. Primer marker size standards (LIZ-120) ar indicated by orange peaks. Fluorescent ddNTPs are colour-coded as follows:- G = Blue, C = Black, T = red and A = green

Table 6.2 Comparison between the primer extension (SnapShot®) and PCR-RFLP analyses.

Sample	Method	$G^{191}A$	$C^{282}T$	$T^{341}C$	$C^{481}T$	$G^{590}A$	$A^{803}G$	$G^{857}A$	*NAT2* Alleles	Genotype
1	RFLP	G/G	C/C		C/T	G/G	G/G	G/G	*2*12A/2*5B*	Intermediate
	ASPCR			T/C						
	PExt.	G/G	C/C	T/C	C/T	G/G	G/G	G/G		
2	RFLP	G/G	C/C		C/C	G/G	A/G	G/G	*2*4/2*5C*	Intermediate
	ASPCR			T/C						
	PExt.	G/G	C/C	T/C	C/C	G/G	A/G	G/G		
3	RFLP	G/G	C/C		C/T	G/G	A/G	G/G	*2*4/2*5B*	Intermediate
	ASPCR			T/C						
	PExt.	G/G/	C/C	T/C	C/T	G/G	A/G	G/G		
4	RFLP	G/G	C/C		C/T	G/G	G/G	G/G	*2*12A/2*5B*	Intermediate
	ASPCR			T/C						
	PExt	G/G	C/C	T/C	C/T	G/G	G/G	G/G		
5	RFLP	G/G	C/T		C/C	G/G	A/G	G/G	*2*4/2*12B*	Fast
	ASPCR			T/T						
	PExt.	G/	C/T	T/T	C/C	G/G	A/G	G/G		
6	RFLP	G/G	C/C		C/T	G/G	G/G	G/G	*2*12A/2*5B*	Intermediate
	ASPCR			T/C						
	PExt	G/G	C/C	T/C	C/T	G/G	G/G	G/G		
7	RFLP	G/G	C/C		C/C	G/G	A/G	G/G	*2*4/2*12A*	Fast
	ASPCR			T/T						
	PExt.	G/G	C/C	T/T	C/C	G/G	A/G	G/G		
8	RFLP	G/G	C/C		C/T	G/G	G/G	G/G	*2*12A/2*5B*	Intermediate
	ASPCR			T/C						
	PExt.	G/G	C/C	T/C	C/T	G/G	G/G	G/G		
9	RFLP	G/G	C/T		C/C	G/A	A/G	G/G	*2*12A/2*6A*	Intermediate
	ASPCR			T/T						
	PExt.	G/G	C/T	T/T	C/C	G/A	A/G	G/G		
10	RFLP	G/G	C/C		T/T	G/G	G/G	G/G		Slow
	ASPCR			C/C					*2*5B/2*5B*	
	PExt.	G/G	C/C	C/C	T/T	G/G	G/G	G/G		
11	RFLP	G/G	C/C		C/T	G/G	G/G	G/G	*2*12A/2*5B*	Intermediate
	ASPCR			T/C						
	PExt.	G/G	C/C	T/C	C/T	G/G	G/G	G/G		
12	RFLP	G/G	C/C		C/C	G/G	G/G	G/G	*2*12A/2*12A*	Fast
	ASPCR			T/T						
	PExt.	G/G	C/C	T/T	C/C	G/G	G/G	G/G		

RFLP = restriction fragment length polymorphism
ASPCR = allele-specific polymerase chain reaction
PExt. = primer extension reaction

NAT2 alleles are classified into their respective haplotype clusters according to the *NAT* website (http://www.louisville.edu/medschool/pharmacology/NAT2.html)

Furthermore, this PE-based technique has also been used with equal success in studies of Alzheimer's disease [364], and for the forensic detection of SNPs in the human mitochondrial genome [365].

We subsequently validated the SNaPshot® technique by genotyping these *NAT2* polymorphisms in an additional 290 TB patients and typical results are depicted in table 6.2. This kit-based method is an efficient and simple technique, which can easily be extended to include other activity-altering *NAT2* SNPs. Several such SNPs, have also been delineated in the human *NAT2* gene, but their purported effect(s) on NAT2 acetylation still need to be validated by *in vitro* analysis. Moreover, there is a dearth of epidemiological information on these SNPs in different populations groups, and therefore, it is difficult to ascertain their significance in the overall acetylation status of individuals. Recently, it was demonstrated that the 364G>A (D122N) and A411T (L137F) SNPs also confer a reduced acetylation activity [180]. However, the significance of these SNPs, as judged by their frequency of occurrence in different population groups has not as yet been reported, to our knowledge. This limitation could easily be addressed using the SNaPshot® technique, by the simple addition of specific primers targeting these SNPs.

6.5 Conclusion

This study describes the use of a simple primer extension methodology which accurately determines the acetylator status in individuals, by characterisation of their *NAT2* genotype. Furthermore, a more encompassing acetylation classification can be achieved by simply including other phenotype-altering SNPs, such as, 364G>A (D122N), 411A>T (L137F), and 434A>C (Q145P), into the overall analysis of *NAT2* acetylation via SNaPshot®. This simple and easy technique enables the high throughput screening of multiple targets in multiple samples, thereby substantially improving the practical application of genotyping in pharmacogenetic studies. Secondly, the genotyping results can easily be converted into a binary code for the purposes of SNP data analyses (1 = homozygous wild-type; 0 = homozygous mutant = 0). This practise could facilitate the automated scoring of the acetylation status.

CHAPTER 7

Allele frequencies for Glutathione *S*–transferase and *N*-acetyltransferase 2 differ in African population groups and may be associated with oesophageal cancer or tuberculosis

Allele frequencies for Glutathione *S*-transferase and *N*-acetyltransferase 2 differ in African population groups and may be associated with oesophageal cancer or tuberculosis incidence.

Craig H. Adams[1*], Cedric J. Werely[1*], Thomas C. Victor[1], Eileen G. Hoal[1], Gawie Rossouw[2] and Paul D. van Helden[1**]

[1]Department of Medical Biochemistry,
[2]Department of Cardiothoracic Surgery,
MRC Centre for Molecular & Cellular Biology, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa

[*]CHA and CJW, equal contribution and joint first authors

[**]E-mail of the corresponding author: pvh@sun.ac.za

Clinical Chemistry and Laboratory Medicine (2003) **41(4):**600-605.

Précis

GST and NAT2 are two drug metabolising enzymes involved in the metabolism of many environmental compounds and chemotherapeutic drugs, and as a result may influence an individual's response to treatment or susceptibility to disease. The Xhosa and SAC ethnic groups of South Africa have a high incidence of Oesophageal Cancer (OC) and Tuberculosis (TB) disease respectively. As a first step towards the case-control investigation of *GST* and *NAT2* polymorphisms as potential markers for OC risk and TB treatment response respectively, we undertook preliminary genotyping analyses of these two genes to establish the baseline frequencies for their allelic variants in healthy individuals from the Xhosa and SAC populations.

We observed a range of frequencies for *GST* and *NAT2* alleles, which may indicate that the metabolic activity of the two enzymes in these groups may be quite different compared to other ethnic groups. For example, we observed a higher frequency with respect to the *NAT2* fast genotype (wild type) in the SAC population than has been reported for Caucasians. *NAT2* fast alleles are believed to reduce the efficacy of the TB treatment regimen due to their rapid clearance of isoniazid from the body. Hence the higher incidence of the fast genotype in SAC individuals could be a factor influencing the response to TB treatment in patients with this ethnic background.

The frequency of the null genotype for *GSTM1* was the lowest thus far reported, whilst the frequency for the *GSTT1* null genotype was considerably higher in the Xhosa group than has been reported for other Africans. The frequency distribution of the *GSTT1* null genotype in the Xhosa population was similar to that reported for Chinese individuals, who incidently also have a very high incidence of OC.

Although further (case-control) studies are required to validate any potential associations between *NAT2* and TB treatment response, or *GSTT1* and OC risk in the populations studied, the allelic frequencies observed for these two genes might be interesting from a clinical perspective, and warrant further investigation.

Sample Source:- Adult samples obtained from the Eastern Cape (Xhosa) and Western Cape Provinces (SAC). Sample cohort used in this study only.

The experimental work was performed by CA (GST) and CJW (NAT2). Both CA and CJW contributed equally to the manuscript preparation. CA and CJW listed alphabetically.

Clin Chem Lab Med 2003; 41(4):600–605 © 2003 by Walter de Gruyter · Berlin · New York

Allele Frequencies for Glutathione *S*-Transferase and *N*-Acetyltransferase 2 Differ in African Population Groups and May Be Associated With Oesophageal Cancer or Tuberculosis Incidence

Craig H. Adams[1]*, Cedric J. Werely[1]*, Thomas C. Victor[1], Eileen G. Hoal[1], Gawie Rossouw[2] and Paul D. van Helden[1]**

[1] Department of Medical Biochemistry,
[2] Department of Cardiothoracic Surgery,
MRC Centre for Molecular & Cellular Biology, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa

Glutathione *S*-transferase (GST) and arylamine *N*-acetyltransferase 2 (NAT2) metabolise many environmental and chemotherapeutic agents, which influence susceptibility to disease. Polymorphisms in these enzymes result in different host phenotypes and contribute to different disease profiles or responses to toxic or chemotherapeutic agents, depending on their frequency in different populations. GST and NAT2 polymorphisms were investigated in different population groups, including African populations, and a range of allelic frequencies have been observed. The *GSTM1* null genotype frequency, reported in this paper in two South African ethnic groups, is the lowest reported (0.19–0.21). In contrast, these same groups have a high *GSTT1* null frequency (0.41–0.54), which is considerably higher than in African-Americans, or other Africans. The *GSTT1* null frequency is comparable to the Chinese, a population with a very high oesophageal cancer incidence, similar to that in the African group. The frequency of the *GSTPi* Val^{105} variant in the South African Xhosas was also high (0.53), differing significantly from the low frequency in other Africans. These variants could therefore be associated with high cancer susceptibility. In addition, the high proportion of *NAT2* "fast" alleles may partially explain the high tuberculosis prevalence in South Africans, due to reduced isoniazid efficacy in the presence of rapid acetylation. Clin Chem Lab Med 2003; 41(4): 600–605

Key words: Glutathione *S*-transferase: GST; Arylamine *N*-acetyltransferase 2: NAT2; African populations.

Abbreviations: COMT, catechol-*O*-methyltransferase; DME, drug metabolising enzyme; GST, glutathione *S*-transferase; INH, isoniazid; NAT2, arylamine *N*-acetyltransferase 2; PCR, polymerase chain reaction; SA, South Africa; TB, tuberculosis.

*CHA and CJW, equal contribution and joint first authors
**E-mail of the corresponding author: pvh@sun.ac.za

Introduction

Any living organism is subject to the insults of chemical attack, whether by therapeutic drugs, carcinogens, or environmental xenobiotics. Different individuals and ethnic groups exhibit varied and sometimes unusual phenotypic responses to these xenobiotic compounds, as has been reported for alcohol metabolism (1, 2), the antihypertensive debrisoquine (3, 4), and isoniazid (INH) (5, 6). Many studies have reported genetic variations in the drug metabolising enzymes (DMEs) to be the cause of these phenotypes (7–9). Consequently, the study of DME polymorphisms has become a major research area, in attempting to understand the genetic causes of individual variability in drug response and susceptibility to toxicity. At the population level, the frequencies and distribution of these alleles are not randomly dispersed, and seem to show ethnic or geographic associations (10–13). Associations between specific allelic variants of some of these enzymes and carcinoma risk have been reported (14–16). Much of the research supports the hypothesis that risk for disease susceptibility may be due to the relative proportions of given allelic variants of some DMEs in different populations.

The classic human acetylation reaction (fast *vs.* slow) has been well studied over the last few decades. The slow acetylation reaction is associated with polymorphic variants of the liver enzyme arylamine *N*-acetyltransferase 2 (NAT2) (17), producing enzymes with variable stability or activity (18). At least 27 different "slow" alleles have been characterised for *NAT2*, but only a select few appear to have an epidemiological impact in different populations (19) (see also the arylamine *N*-acetyltransferase webpage: http://www.louisville.edu/medschool/pharmacology/NAT.html/). INH is primarily metabolised by NAT2, and the enzyme activity is dependant upon the specific *NAT2* alleles present in individuals, and hence, the population. Therefore, given that the incidence of tuberculosis (TB) is increasing world-wide (and assuming pandemic proportions in certain populations) it may be expedient to investigate NAT2 polymorphisms in populations with a high TB incidence. Certain forms of NAT2 associated with populations with a high TB incidence, may be responsible for an increased risk of developing drug resistance, failure to cure, or adverse side effects.

The glutathione *S*-transferases (GST) are a family of DMEs that play an important role in the phase II detoxification of a number of carcinogens. They are also involved in the regulation of other enzymes that in turn are involved in the maintenance of the genomic integrity of the cell (20). In addition to its detoxification function,

GSTT1 also has the ability to activate certain carcinogens, causing tumours in mice (21). As such, GSTs may contribute indirectly to cancer susceptibility and predisposition. Of the five classes of GST enzymes, the GSTM1 and GSTT1 enzymes have been studied extensively in relation to the aetiology of various cancers, particularly since the deletion variants of these enzymes (the null genotypes) abrogate the activity of that enzyme in the individual. These individuals are thus unable to metabolise certain xenobiotics or reactive intermediates (22), leading to accumulation of DNA adducts in the cell, and subsequent carcinogenesis. Moreover, the expression of GSTT1 is not only restricted to the liver but also occurs in red blood cells, affording this enzyme a greater role in the body's detoxification reaction systems than GSTM1. Hence the null genotype of this enzyme may have a larger impact on the body's ability to detoxify xenobiotic compounds. The frequencies of the null genotypes of *GSTM1* and *GSTT1* show substantial variation in different population groups (23) and may therefore affect the risk of cancer in these groups.

GSTPi is a major enzyme involved in the detoxification of cigarette smoke carcinogens (20). Two polymorphisms at codons 105 and 114, resulting in amino acid changes ($I^{105}\rightarrow V$ and $A^{114}\rightarrow V$), have been described for GSTPi (24). These polymorphisms have been reported to be strongly associated with bladder and testicular cancer (25), and may be important in other forms of cancer, such as oesophageal cancer.

Africans currently comprise approximately 20% of the global population, and recent genetic analyses suggest they are the most genetically diverse group of people on earth (26, 27). Therefore, we believe it is necessary to study many groups of Africans (classified according to language, for example) and not assume that all Africans will have similar phenotype and genotype characteristics. In order to test this hypothesis, we studied polymorphisms in the *GST* and *NAT2* genes in two major ethnic groups from South Africa and compared these to other population groups. This baseline data may be helpful in determining policy guidelines for cancer prevention and chemotherapy in African, as well as other populations.

Materials and Methods

Study subjects

Individuals from the Coloured and Xhosa population groups were enrolled in this study after informed consent and ethical approval of the project by the institutional Ethics Committee of Stellenbosch University. All Xhosa subjects were Xhosa-speaking, black Africans, originating from the eastern Cape region of South Africa (SA). The Coloured subjects were representative of the Coloured population, resident in the western Cape region of South Africa. This group comprises a large and potentially genetically diverse population, supporting genetic contributions from Caucasoid, Khoisan, Malay, and Negroid ancestry (28). It is the largest of the ethnically distinct subgroups that inhabit the western Cape region, comprising approximately 10% of the population of SA, and will be referred to as Coloured throughout this study. All individuals were apparently healthy at the time of sampling.

Genotyping for GSTM1, GSTT1, GSTP1 and NAT2

GSTM1 and T1 polymorphisms were assayed *via* a modified, multiplex PCR protocol as described by Bailey *et al* (29). The amplification reaction was carried out in a volume of 100 µl consisting of 400 ng of genomic DNA, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 1.5 U Taq DNA Polymerase I (Promega Corp., Madison, Wisconsin, USA), 0.5 µM primers (being M1 [5′-CTGCCTACTTGATTGATGGG-3′], and M2 [5′-CTGGATTGTAGCAGATCATGC-3′] for the specific amplification of *GSTM1*, and primers T1 [5′-TTCCTTACTGGTCCTCACATCTC-3′], and T2 [5′-TCACCGGATCATG GCCAGCA-3′] for specific *GSTT1* sequence amplification). This reaction mixture also contained 0.5 µM of a third primer set ($COMT^{For}$, [5′-TCACCATCGAGATCAACCCC-3′], and $COMT^{Rev}$, [5′-ACAACGGGTCAGGCAT GCA-3′]) for the specific amplification of a fragment of the catechol-*O*-methyltransferase (*COMT*) gene (30), to serve as internal control. The amplification conditions consisted of an initial denaturation step (95 °C for 2 min), followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 64 °C for 60 s, and polymerization at 72 °C for 120 s. The PCR results were analysed by polyacrylamide gel electrophoresis (PAGE), in a Mini Protean II gel apparatus (Biorad laboratories, Hercules, USA) using 12% polyacrylamide. The profile of bands was visualised by subsequent ethidium bromide staining of the gels. *GSTT1* null genotypes were assayed by the absence of a 480 bp DNA band, whilst the absence of a band at 253 bp determined *GSTM1* null genotypes. The presence of the COMT-specific 95 bp band eliminated the possibility of erroneously scoring null genotypes as a result of PCR failure.

The *GSTP1* codon 105 polymorphism was assessed by PCR amplification of an 851 bp *GSTP1*-specific DNA band, as described by Harris (31). The amplification was carried out in a 50 µl reaction volume, containing 200 ng genomic DNA, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 1.5 U Taq DNA polymerase I, and 0.5 µM primers P1 [5′-CGCATGCTGCTGGCAGATCAG-3′] and P2 [5′-CAAGCCACCTGAGGGG TAAGG-3′]. Thermal cycling was carried out under the following conditions: initial denaturation at 95 °C for 5 min, followed by 30 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s. The 851 bp product was subsequently restricted with *Bsm*A1 (New England Biolabs, Beverly, MA, USA), and the results assessed by 12% PAGE as above. A similar genotyping strategy was also followed to detect the GSTP1 polymorphism at codon 114, using 0.4 µM of primers P3 [5′-GTTGTGGGGAGCAAGCAGAGG-3′] and P4 [5′-CACAATGAAGGTCTTGCCT CCC-3′], and the following cycling conditions: 95 °C for 5 min, followed by 30 cycles at 94 °C for 30 s, 62 °C for 30 s, and 72 °C for 30 s. The amplified 217 bp DNA band was then subjected to *Aci*I (New England Biolabs) endonuclease restriction, to type the allelic variation at this codon. Gel analysis was by 12% PAGE as described above.

The analysis of the NAT2 polymorphisms proceeded by the PCR-based genotyping strategy as previously described (32) Briefly this entailed the PCR amplification of a 1000 bp fragment of the *NAT2* gene in a 100 µl reaction volume, followed by cleavage of separate PCR aliquots with *Kpn*I, *Bam*HI, *Msp*I, and *Taq*I, respectively. DNA profiles were then analysed by horizontal gel electrophoresis using Metaphor® agarose (BMA, Rockland, Maine, USA) and visualised with GelStar® nucleic acid gel stain (BMA, Rockland, Maine, USA) according to the manufacturer's recommendations.

Results

The *GSTM1* and *GSTT1* null genotype frequencies in the SA cohorts (including Xhosa and Coloured) in com-

parison to other ethnic groups are shown in Table 1. The *GSTM1* null frequency for Caucasians and Chinese is notably high and ranges between (0.40–0.60). In contrast, the frequency in African-Americans and the other African ethnic groups (including Zimbabweans, Venda, and Tanzanians) is considerably lower (0.28–0.35). In the SA cohort (combined n = 197), a similar low frequency (0.20) was observed and this differed significantly (χ^2=75.84; p < 0.0001) from that of the Chinese and Caucasian populations.

A varied distribution of the *GSTT1* null genotype was also evident among ethnic groups. A relatively low *GSTT1* null frequency (0.20–0.26) was observed for African-Americans and the other African (Zimbabwe, Venda, and Tanzanians) ethnic groups. This low *GSTT1* null frequency complements the low *GSTM1* null frequency observed. A similar low *GSTT1* null frequency (0.10–0.24) is observed for Caucasians, who in contrast, have a high *GSTM1* null frequency. The Chinese population, which has a high *GSTM1* null frequency, also has a relatively high *GSTT1* null frequency (0.64). Similarly, a high *GSTT1* null frequency (0.40–0.57) was observed for the SA Xhosa and Coloured ethnic groups, which is significantly different (χ^2=27.24; p < 0.0001) from that observed for the other African ethnic groups. A greater than two-fold higher frequency of the double null genotype was also evident among the SA cohort, compared to the other African ethnic groups.

Table 2 shows the genotype distribution of the GSTP1 polymorphism at codon 105 for SA Xhosa compared to various other ethnic groups. The *GSTP1* Ile^{105} to Val^{105} change was observed at an allelic frequency of 0.53 for SA Xhosa, which is significantly greater than that reported for Tanzanians (χ^2=62.67; p < 0.0001), Vendas (χ^2=69.75; p < 0.0001), and Zimbabweans (χ^2=34.40; p < 0.0001). This allelic frequency is similar to that observed for African-Americans but contrasts the low allelic frequency (0.14–0.18) observed for Asians. The analysis of the GSTPi polymorphism at codon 114 showed that the V^{114} allele was absent from the Xhosa population (results not shown).

The *NAT2* slow acetylator allele frequency in diverse populations has been compared in Tables 3 and 4. Cer-

Table 1 Comparison of the null genotype frequency for *GSTM1* and *GSTT1* in the South African cohorts and other ethnic groups.

Population group	GST null genotypes		
	*GSTM1*0* (%)	*GSTT1*0* (%)	*GSTM1*0/GSTT1*0* (%)
African-Americans (33)	28–35	22	NA
Caucasians	40–60	10–24	NA
Chinese (33–35)	40–60	64	NA
Zimbabweans (36)	24	26	5.7
Vendas (36)	23	20	5.4
Tanzanians (36)	33	25	4.2
South African Xhosa (n = 128)	21	41	11
South African Coloured (n = 69)	20	57	13

NA = not available.

Table 3 Population frequency of the fast acetylators for arylamine *N*-acetyltransferase 2.

Population group	Fast genotype	Fast phenotype
Indians (41)	0.26	≈0.2
Caucasians (USA) (41)	0.24	0.25
African-Americans (41)	0.41–0.32	0.41
Japanese, Koreans (41)	0.74–0.64	0.66
SA Coloured (n = 170)	0.56	ND
SA Xhosa (n = 240)	0.60	ND

ND = not determined.

Table 2 Distribution of the *GSTP1*-codon 105 polymorphism (Ile^{105}→Val^{105}).

Population group	*GSTP1* Ile^{105}/Val^{105} genotype			
	% Ile/Ile (No./total)	% Ile/Val (No./total)	% Val/Val (No./total)	% frequency of Val^{105} allele (No./total)
Tanzanian (36)	77% (79/102)	14% (14/102)	9% (9/102)	16% (32/204)
Venda (37)	77% (67/86)	22% (19/87)	1% (1/87)	14% (21/174)
Zimbabwean (37)	58% (56/97)	36% (35/97)	6% (6/97)	24% (47/194)
Caucasian (25, 38)				28–32%
African-American (39)				42%
Asian (40)				14–18%
Indian (40)				27%
SA Xhosa	22% (22/101)	51% (51/101)	28% (28/101)	53% (107/202)

Table 4 Frequency of the *NAT2* slow acetylator alleles in different population groups.

Population group	No. of alleles	Frequency of the slow acetylator alleles				
		481C→T (*2*5*)	590G→A (*2*6*)	857G→A (*2*7*)	191G→A (*2*14*)	Combined *2*5 to 2*14*
US White (41)	98	0.41	0.34	0.020	0	0.76
African-Americans (41)	172	0.29	0.24	0.07	0.076	0.68
Indians (41)	122	0.33	0.38	0.033	0	0.74
Amerindians (42)	482	0.024–0.1	0–0.037	0.23	ND	0.33
Japanese (41)	158	0.019	0.23	0.11	0	0.36
Koreans (41)	170	0.018	0.18	0.11	0.011	0.32
Hong Kong (41)	140	0.057	0.31	0.16	0	0.52
Chinese (43)	168	0.071	0.27	0.053	–	0.41 (ref. 44)
Taiwanese (41)	200	0.025	0.31	0.15	0	0.48
Filipinos (41)	200	0.065	0.36	0.18	0	0.60
Samoans (41)	50	0.040	0.34	0.02	0	0.40
Asians (45)						0.14
Finns (44)						0.76
Swedes (44)						0.82
Danes (46)						0.54
SA Coloured	170	0.25	0.11	0.029	0.047	0.44
SA Xhosa	240	0.20	0.15	0.017	0.025	0.39

ND = not determined.

tain alleles, *e.g.*, *NAT2*14*, are absent from most of the population groups represented here, whilst others *e.g.*, the *2*5* allele, have a very low frequency in the Oriental groups. The *2*7* allele is the most abundant in the Amerindians, whereas this allele has an incidence of less than 1% in all of the other groups. Significant differences in the *2*7* allele frequency are also evident between these various population groups. It is also notable that the combined slow allele frequency in African-Americans is 0.68, compared to 0.39–0.44 in the two African populations in our study. The frequencies of specific alleles also show significant differences as detailed in Table 4.

Discussion

GSTs are involved in the detoxification of a wide variety of chemical carcinogens, and four main subclasses have been identified in humans (20) of which three are known to be polymorphic. In this study, the frequency of the homozygous deletion of *GSTM1* and *GSTT1* (null genotype) was determined for a SA cohort and these were compared to other ethnic groups. The *GSTM1* null frequency was low in the SA cohort at 0.20, which is in agreement with the range of between 0.28 and 0.35 that has been reported for other African populations (34, 35). This, however, is in contrast to the high frequency of between 0.4 and 0.6 that has been reported for Caucasians (47) and Chinese (33–35).

The *GSTT1* null frequencies were observed to be highly variable among populations, significantly so for the SA cohort (0.40–0.57) compared to that reported (0.20–0.26) for other African ethnic groups (36, 37). Interestingly, a similar high *GSTT1* null frequency is evident in the Chinese population (0.64), which, like the SA Xhosas, has a high incidence of oesophageal cancer.

The third subclass of GST known to be polymorphic is GSTP1. There are two single nucleotide polymorphism (SNPs) in the coding region of GSTP1 that have been characterised, resulting in amino acid changes of Ile^{105} to Val^{105} in one instance, and Ala^{114} to Val^{114} in the other. The *GSTP1* Val^{105} variant has been associated with low substrate affinity and thus reduced enzyme activity (25). The allelic frequency of the *GSTP1* Val^{105} variant was determined for the SA Xhosa population and was noted to be significantly high in comparison to that reported for other African ethnic groups. Since this variant is associated with reduced activity, a high frequency could increase the risk of smoking-induced diseases *e.g.*, oesophageal cancer, due to impaired detoxification.

From a quantitative point of view, both with respect to the number of individuals affected, as well as the number of different compounds metabolised, NAT2 can be ranked as one of the most important DMEs. Furthermore, several of the important alleles (such as *2*5*, *2*6*, *2*7*, and/or *2*14*) – measured by frequency of occurrence in populations – have a number of specific mutations, which are characteristic of that allele. As has been suggested by Meisel *et al.* (48), this combination of mutations for the separate alleles may contribute to a continuum of varying metabolic phenotypes, which may be more pronounced in some ethnic groups rather than others, as can be seen from the comparisons in Table 4. This could account for the difference in the phenotypic distribution of NAT2 activity that has been described for some population groups, such as bimodal in Caucasians (6), and trimodal in Japanese (5), and in SA Coloured (32).

The type of mutant *NAT2* allele(s) inherent in the in-

dividual may also be significant in defining the NAT2 acetylation reaction. Secondly, the hypothesis of ethnically linked *NAT2* alleles, as has been proposed for the *2*14* allele (the so called African-allele), and which has been reported in African-Americans, is supported by our findings of this allele in our population cohorts. However, the *NAT2*14* allele has also been reported in Hispanics and Koreans resident in the USA (41).

In vitro studies (18) using expressed variants of the *NAT2*5*, *2*6*, *2*7*, and *2*14* alleles indicate that the activity of these proteins varies greatly over a period of 6 hours (as compared to the wild-type *NAT2*4*). The *NAT2*5* variant maintained approximately 75% of its activity, whereas the other mutant NAT2 proteins had lost almost all activity over this period. This type of result serves to underline the hypothesis that the prevalence of various mutant *NAT2* alleles in a population would have a significant effect on acetylation. Hence it is important to determine the *NAT2* alleles in diverse ethnic groups and populations, particularly in regard to the pandemic incidence of tuberculosis in some geographic locations in the world. The efficacy of antituberculosis chemotherapy and/or preventative measures may therefore require optimisation for various ethnic groups in the light of their inherent NAT2 status.

Acknowledgements

We gratefully acknowledge the assistance of Mr. F.A. Venter, Dr. W.C.A. Gelderblom (Medical Research Council Program on Mycotoxins & Experimental Carcinogenesis, P.O. Box 19070 Tygerberg), and the funding by the Cancer Association of South Africa (CANSA).

References

1. Vesell ES, Page JG, Passananti GT. Genetic and environmental factors affecting ethanol metabolism in man. Clin Pharmacol Ther 1971; 12:192–201.
2. Reed TE. Racial comparisons of alcohol metabolism: background, problems and results. Alcoholism: Clin Exp Res 1978; 2:83–7.
3. Mahgoup A, Idle JR, Dring LG, Lancester R, Smith RL. Polymorphic hydroxylation of debrisoquine in man. Lancet 1977; 2:584–6.
4. Eichelbaum M, Spannbrucker N, Steincke B, Dengler JJ. Defective N-oxidation of sparteine in man: a new pharmacogenetic defect. Eur J Clin Pharmacol 1979; 16:183–7.
5. Sunahara S, Urano M, Ogawa M. Genetic and geographic studies on isoniazid inactivation. Science 1961; 134:1530.
6. Evans DAP, White TA. Human acetylation polymorphism. J Lab Clin Med 1964; 63: 394–403.
7. Grant DM, Blum M, Beer M, Meyer UA. Monomorphic and polymorphic human arylamine N-acetyltransferases: a comparison of liver isoenzymes and expressed products of two cloned genes. Mol Pharm 1991; 39:184–91.
8. Smith CAD, Smith G, Wolf CR. Genetic polymorphisms in xenobiotic metabolism. Eur J Cancer 1994; 30A:1921–35.
9. Zhong S, Howie AF, Ketterer B, Taylor J, Hayes JD, Beckett GJ, *et al.* Glutathione S-transferase mu locus: use of genotyping and phenotyping assays to assess association with lung cancer susceptibility. Carcinogenesis 1991; 12:1533–7.
10. Taioli E. International collaborative study on genetic susceptibility to environmental carcinogens. Cancer Epidemiol Biomark Prev 1999; 8:727–8.
11. Garte S. The role of ethnicity in cancer susceptibility gene polymorphisms: the example of CYP1A1. Carcinogenesis 1998; 19:1329–32.
12. Watson MA, Stewart RK, Smith GB, Massey TE, Bell DA. Human glutathione *S*-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. Carcinogenesis 1998; 19:275–80.
13. Weber WW. Populations and genetic polymorphisms. Molec Diagnosis 1999; 4:299–307.
14. Seidegard J, Pero RW, Miller DW, Beattle EJ. Glutathione transferase in human leucocytes as a marker for the susceptibility to lung cancer. Carcinogenesis 1986; 7:751–3.
15. Zhong S, Wyllie AM, Barnes D, Wolf CR, Spurr NK. Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer. Carcinogenesis 1993; 14:1821–4.
16. Risch A, Wallace DM, Bathers S, Sim E. Slow *N*-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer. Hum Mol Genet 1995; 4:231–6.
17. Hein DW, Doll MA, Rustan TD, Ferguson RJ. Metabolic activation of N-hydroxyarylamines and *N*-hydroxyarylamides by 16 recombinant human NAT2 allozymes: effects of 7 specific NAT2 nucleic acid substitutions. Cancer Res 1995; 55:3531–6.
18. Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A, *et al.* Human acetyltransferase polymorphisms. Mut Res 1997; 376:61–70.
19. Hein DW, Doll MA, Fretland AJ, Leff MA, Webb SJ, Xiao GH, *et al.* Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. Cancer Epidemiol Biomark Prev 2000; 9:29–42.
20. Hayes J, Pulford D. The glutathione *S*-transferase supergene family: regulation of GST* and the contribution of the isoenzymes to cancer chemotherapy and drug resistance. Crit Rev Biochem Mol Biol 1995; 30:445–600.
21. Landi S. Mammalian class theta GST and differential susceptibility to carcinogenesis: a review. Mutation Res 2000; 463:247–83.
22. Vineis P, Malats N, Boffetta P. Why study metabolic susceptibility to cancer? IARC Sci Publ 1999; 148:1–3.
23. Geisler SA, Olshan AF. GSTM1, GSTT1 and risk of squamous cell carcinoma of the head and neck. Am J Epidemiol 2001; 154:95–105.
24. Board PG, Webb GC, Coggan M. Isolation of a cDNA clone and localization of the human glutathione S-transferase 3 gene to chromosome bands 11q13 and 12q13-14. Ann Hum Genet 1989; 53:205–13.
25. Harries LW, Stubbins MJ, Forman D, Howard GC, Wolf CR. Identification of genetic polymorphisms at the glutathione S-transferase Pi locus and association with susceptibility to bladder, testicular and prostate cancer. Carcinogenesis 1997; 18:641–4.
26. Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA, *et al.* Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am J Hum Gen 1995; 57: 523–38.
27. Melton T, Ginther C, Sensabaugh G, Soodyall H, Stoneking M. Extent of heterogeneity in mitochondrial DNA of subsaharan African populations. J Forensic Sci 1997; 42:582–92.
28. Marais JS. The Cape Coloured People: 1652 to 1932. 1st ed. Johannesburg: Witwatersrand University Press, 1957.
29. Bailey LR, Roodi N, Verrier CS, Yee CJ, Dupont WD, Parl FF.

Breast cancer and CYP1A1, GSTM1 and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African-Americans Cancer Res 1998; 58:65–70.
30. Karayiorgou M, Altemus M, Galke BL, Goldman D, Murphy DL. Genotype determining low catechol-O-methyltransferase activity as a risk factor for obsessive-compulsive disorder. Proc Natl Acad Sci USA 1997; 94:4572–5.
31. Harris CC. Chemical and physical carcinogenesis: advances and prospectives for the 1990s. Cancer Res 1991; 51:5023–44.
32. Parkin DP, Vandenplas S, Botha FJH, Vandenplas M, Seifart HI, van Helden PD, *et al.* Trimodality of isoniazid elimination: phenotype and genotype in patients with tuberculosis. Am J Crit Care Med 1997; 155:1717–22.
33. Chen CL, Liu Q, Relling MV. Simultaneous characterization of glutathione S-transferase Mi and T1 polymorphism by polymerase chain reaction in American whites and blacks. Pharmacogenetics 1996; 6:187–91.
34. Baranova H, Perriot J, Albuisson E, Ivaschenko T, Baranov, VS, Hemery B, *et al.* Pecularities of the GSTM1 0/0 genotype in French heavy smokers with various types of chronic bronchitis. Hum Genet 1997; 99:822–6.
35. Bruhn C, Brockmoller J, Kerb R, Roots I, Borchert H. Concordance between enzyme activity and genotype of glutathione S-transferase theta (GSTT1). Biochem Pharmacol 1998; 56:1189–93.
36. Dandara C, Sayi J, Masimirembwa CM, Magimba A, Kaayi S, *et al.* Genetic polymorphism of cytochrome P450 1A1 (CYP1A1) and glutathione transferases (M1, T1 and P1) among Africans. Clin Chem Lab Med 2002; 40:1–6.
37. Masimirembwa CM, Dandara C, Sommers KD, Snyman JR, Hasler JA. Genetic polymorphism of cytochrome P4501, microsomal epoxide hydrolase and glutathione S-transferase M1 and T1 in Zimbabweans and Venda of South Africa. Pharmacogenetics 1998; 8:83–5.
38. Landi S, Norppa H, Frenzilli G, Cipolllini G, Ponzanelli I, Barale R, *et al.* Individual sensitivity to cytogenetic effects of 1,2:3,4-diepoxybutane in cultured human lymphocytes: influence of glutathione S-transferae M1, P1 and T1 genotypes. Pharmacogenetics 1998; 8:461–71.
39. Watson MA, Stewart RK, Smith GB, Massey TE, Bell DA. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue activity and population frequency distribution. Carcinogenesis 1998; 19:275–80.
40. Katoh T, Kaneko S, Takasawa S, Nagata N, Inatomi H, Ikemura K, *et al.* Human glutathione S-transferase P1 polymorphism and susceptibility to smoking related epithelial cancer, oral lung, gastric, colorectal and urothelial cancer. Pharmacogenetics 1999; 9:165–9.
41. Lin HJ, Han C-Y, Hardy S. Ethnic distribution of slow acetylator mutations in the polymorphic N-acetyltransferase (NAT2) gene. Pharmacogenetics 1994; 4:125–43.
42. Jorge-Nebert LF, Eichelbaum M, Griese E-U, Inaba T, Arias TD. Analysis of six SNPs of NAT2 in Ngawbe and Embera Amerindians of Panama and determination of the Embera acetylation phenotype using caffeine. Pharmacogenetics 2002; 12:39–48.
43. Ellard GA. Variations between individuals and populations in the acetylation of isoniazid and its significance for the treatment of pulmonary tuberculosis. Clin Pharmacol Ther 1976; 19:610–25.
44. Weber WW. The acetylator genes and drug response. New York: Oxford University Press, 1987.
45. Yu MC, Skipper PL, Taghizadeh K, Tannenbaum SR, Chan KK, Henderson BE, *et al.* Acetylator phenotype, aminobiphenyl-haemoglobin adduct levels, and bladder cancer risk in white, black and Asian men in Los Angeles, California. J Natl Cancer Inst 1994; 86:712–16.
46. Hein DW. The acetylator genotype and arylamine-induced carcinogenesis. Biochim Biophys Acta 1988; 948:37–66.
47. Garte S, Gaspari L, Alexandrie A-K, Ambrosone C, Autrup H, Autrup JL, *et al.* Metabolic gene polymorphism frequencies in control populations. Cancer Epidemiol Biomark Prev 2001; 10:1239–48.
48. Meisel P, Arndt D, Scheuch E, Klebingat K-J, Siegmund W. Prediction of metabolic activity from genotype: the gene-dose effect of N-acetyltransferase. Ther Drug Monit 2001; 23:9–14.

Received 3 December 2002, revised 24 January 2003, accepted 30 January 2003

Corresponding author: Prof. Paul D. van Helden, Department of Medical Biochemistry, Stellenbosch University, P. O. Box 19063, Tygerberg 7505, South Africa
Phone: +27 21-9389401, Fax: 2721-9317841,
E-mail: pvh@sun.ac.za

CHAPTER 8

NAT2 slow acetylator function as a risk indicator for age-related cataract formation

NAT2 slow acetylator function as a risk indicator for age-related cataract formation

David Meyer[a], Donald P. Parkin[b], Heine I Seifart[b], Stefan Maritz[d], Albertus H. Engelbrecht[a], Cedric J. Werely[c], and Paul van Helden[c]

Departments of [a]Ophthalmology, [b]Pharmacology, and [c]Medical Biochemistry, Faculty of Health Sciences, University of Stellenbosch, P.O. Box 19063, Tygerberg 7505, South Africa, and [d]Department of Biostatistics of the South African Medical Research Council. P.O. Box 19070, Tygerberg 7505, South Africa.

Pharmacogenetics (2003) **13:**285-289.

Précis

Cataract disease is a progressive, degenerative ocular condition, and various environmental and genetic factors, as well as certain clinical conditions have been implicated in cataract formation. Advancing age appears to be the most significant factor associated with the disease globally. We therefore hypothesized that increasing levels of toxicants from environmental insults, coupled to a reduced metabolic turnover of these substances may contribute to cataract formation.

We investigated the NAT2 acetylation status in cataract disease in SAC individuals in a case control association study. We observed complete concordance between the NAT2 phenotype and genotype results. Furthermore, the *NAT2* Slow allele (S) was significantly associated with cases rather than control samples, which indicated that the slow allele was a significant risk factor in cataract disease. Similarly a comparison of the fast (FF), intermediate (FS), and slow (SS) acetylators also showed that the greatest risk for cataract disease was associated with individuals having the Slow (SS) genotype. Individuals having the intermediate genotype are also at risk for cataract disease, albeit at a lower risk level than for the Slow genotype.

Unfortunately, NAT2 acetylation activity has not as yet been reported inside the human eye, and therefore the acetylation activity in loco is uncertain at this stage. However, it is interesting to note that acetylation activity has been found in the bovine lens as well as in the rabbit eye. These results indicate that there may well be a role for NAT acetylation in the eye. Further studies investigating the contribution(s) of the NAT enzymes inside the eye would be a valuable addition to the already published work.

Sample Source:- Samples were case-control matched SAC adults from the Western Cape Province. Sample cohort used in this study only.

All the NAT2 genotypic analysis was performed by CJW and the original draft manuscript was prepared by CJW and PDvH.

NAT2 slow acetylator function as a risk indicator for age-related cataract formation

David Meyer[a], Donald P. Parkin[b], Heine I. Seifart[b], J. Stefan Maritz[d], Albertus H. Engelbrecht[a], Cedric J. Werely[c] and Paul D. van Helden[d]

Objectives **To show that the slow arylamine *N*-acetyltransferase type 2 (NAT2) catalysed acetylator function is associated with the development of age-related cataracts.**

Methods **Both the acetylator phenotype and genotype of 139 patients with age-related cataracts were determined, and the distribution of the acetylator subtypes in the cased population was compared with the distribution in the general (control) population. The genotype was determined by restriction-enzyme analysis of DNA, and the phenotype was determined using the elimination characteristics of isoniazid as discriminant.**

Results **The frequency of alleles coding for slow acetylator characteristics was higher in the patients than in the controls, and the difference was significant ($P = 0.013$).**

Conclusions **Slow acetylators are at higher risk of developing age-related cataracts than fast acetylators and we suggest that exogenous factors, which can be detoxified by acetylation, are aetiological agents for cataract formation. Identification of and avoidance of such (environmental) agents should reduce the incidence of age-related cataracts. *Pharmacogenetics* 13:285–289**

Pharmacogenetics 2003, 13:285–289

Keywords: cataracts, age, acetylator status

Departments of [a]Ophthalmology, [b]Pharmacology and [c]Medical Biochemistry, Faculty of Health Sciences, University of Stellenbosch, PO Box 19063, Tygerberg 7505, South Africa and [d]Department of Biostatistics of the South African Medical Research Council, PO Box 19070, Tygerberg 7505, South Africa.

Correspondence to D. Meyer, Department of Ophthalmology, PO Box 19063, Tygerberg, 7505, South Africa.
Tel: +27 21 9389380; fax: +27 21 9318437; e-mail: dm2@sun.ac.za

Received 19 November 2002
Accepted 12 February 2003

Introduction

Humans are constantly exposed to foreign chemicals, both man-made and natural. These foreign chemicals, or xenobiotics, may be drugs, industrial chemicals, pesticides, or toxins from plants, animals and moulds. Removal of these compounds from the body often relies on enzymatic conversion (i.e. metabolism) in the liver and other tissues to more readily excitable water-soluble forms.

Individuals differ in their ability to metabolize drugs and other xenobiotics due to inherited variations in the gene sequence of their biotransforming enzymes. These genetic differences called polymorphisms, can increase or decrease drug metabolism and in some cases, even eliminate a metabolic pathway. Such polymorphisms have been associated with a wide array of clinical conditions. One such known polymorphic enzyme is arylamine *N*-acetyltransferase 2, (NAT2; E.C. 2.3.1.5), which can acetylate a variety of arylamine compounds. The NAT2 gene contains a number of sites that are polymorphic [1], and these allelic variants have been reported to be co-dominant [2]. The alleles encode a fast (F) or slow (S) enzyme variant and individuals may be phenotypically fast (FF), intermediate (FS) or slow acetylators (SS), depending on which two alleles are present. Slow acetylator status has been associated with conditions such as urinary bladder cancer [3–8], epilepsy [9], Gilbert's syndrome [10], endometriosis [11], renal cell carcinoma [12], malignant mesothelioma [13,14], oesophageal carcinoma [15], asthma [16], allergic and atopic disease [17] and autoimmune diseases such as systemic lupus erythematosus and systemic sclerosis [18]. Rapid acetylator status on the other hand has been associated with conditions such as colorectal cancer [19–21] type 1 diabetes mellitus [22], lung cancer [23], glomerulonephritis [24], benign breast disease [25], laryngeal cancer [26] and phenylketonuria [27].

However, no clinical association has, to our knowledge, been made between degenerative ocular conditions such as cataract and age-related macular degeneration on the one hand, and genetic polymorphisms on the other. Cataracts in humans occur worldwide, but especially in developing countries. They are one of the major causes of visual impairment, eventually leading to blindness, if left un-operated. It is now well established that the aetiology of human cataract formation is multifactorial [28,29]. This is apparent from the association of cataracts with a number of clinical conditions. Cataract can also be induced by nutritional deficiency

0960-314X © 2003 Lippincott Williams & Wilkins

DOI: 10.1097/01.fpc.0000054077.64000.3a

of certain amino acids (such as tryptophan), proteins (hypoproteinemia) and hypovitaminosis-B_2 [28]. Exposure to environmental factors such as various wavelengths of ionizing as well as non-ionizing radiations, heavy metals, sunlight, cigarette smoke, oxygen and oxy-radicals have all been associated with cataract formation [29]. It is most likely that cataract results from an interaction between genetic and environmental factors; hence the observation that age is probably recognized as the most significant factor of association amongst patients in the above 50-year-old age group.

N-acetylation is involved in a wide variety of detoxification processes of, *inter alia*, environmental chemicals, and hence, this study was conducted to examine the relationship between the acetylator pheno- and genotypes in a group of patients with mature age-related cataract in a well-defined segment of the population of the Western Cape region of South Africa, namely, the coloured subgroup. The coloured (mixed race) population is the largest of the ethnically distinct subgroups that inhabit the region. This South African population of mixed ancestry includes Malay, Khoisan, Negroid and Caucasoid stock, and is referred to as coloured throughout this study [30].

Distribution of polymorphic *N*-acetyltransferase genotypes and phenotypes in this population has previously been well documented by us [31,32]. Hence this well studied and special population group was selected for the purpose of this study.

Methods

Patients and controls

One hundred and thirty-nine adult patients of both genders aged 50 and above, with classic mature age-related cataract, presenting for lens extraction and intra-ocular lens implantation surgery, were enrolled in the study, after informed consent and for which approval was obtained from the institutional review board of the University of Stellenbosch. Patients were included in the study if they perceived themselves to be coloured (of mixed race) and if this was verified by at least one independent observer. Care was taken to exclude all patients with well-recognized aetiological factors for cataract formation, such as diabetes mellitus, previous ocular trauma, metabolic and/or inherited diseases. The patient sample therefore consisted of cases with age-related cataracts of unknown aetiology.

The coloured population of the region was used as the reference population, against which to assess deviation of the distribution of acetylator genotypes/phenotypes in the trial group from the normal, since all participants in the trial were drawn from this ethnically distinct and close-knit community. The distribution of acetylator phenotypes/genotypes in this population has previously been determined in our laboratory [31].

A patient was diagnosed to have an age-related cataract if he or she was at least 50 years of age, had no chronic intercurrent condition known to predispose to cataractogenesis (e.g. diabetes mellitus, hyperlipidaemia, long-term treatment with glucocorticosteroids), and had not previously suffered trauma to the affected eye [22, 33–36].

The age of the patients varied from 50.16 to 86.58 years (mean = 67.4 years), and the group comprised 60% (83/139) females and 40% (56/139) males. The mean body mass was 63.15 kg.

Acetylator phenotype

All patients participating in the trial were admitted to hospital for pre-operative evaluation one day prior to lens implant surgery. On this day an oral test dose of isoniazid (10 mg/kg) was administered one hour before a light breakfast at 07:00 hours. Four blood samples (2 ml each) were then collected from the antecubital vein, at the following times after the dose: 2, 3, 4 and 5 h. The isoniazid concentration in plasma was determined by high-pressure liquid chromatography [37]. The first-order elimination rate constant was determined from the concentration vs. time data and was used as the phenotype parameter. Patients were classified as slow, intermediate or fast acetylator phenotypes using the criteria of Parkin *et al.* [31].

Acetylator genotype

A 5-ml blood sample was collected in ethylenediaminetetra-acetic acid tubes from each patient for the determination of the NAT2 genotype profile as previously described [31]. The genomic DNA was extracted using a salting out procedure [38] and a 1000 bp sequence coding for the NAT2 enzyme was amplified by PCR, as previously described [31]. The 1000 bp segment was then analysed using the restriction fragment length polymorphism technique, making use of four enzymes for appropriate cleavage (*Bam*H1, *Kpn*I, *Mp*I and *Taq*I). The NAT2 DNA cleavage profiles generated by each of the restriction enzymes were analysed by gel electrophoresis using MetaPhor® agarose (Biowhittaker, USA), and the profiles were visualized under ethidium bromide staining.

Results

The study population consisted of 139 consecutive cases (50 years and older) with mature age-related cataract with absolute concordance in a trimodal pattern between the isoniazid (INH) eliminator phenotype and the NAT2 genotype, compared to 114 controls meeting the same criteria.

The cataract group represented the typical age-related cataract population as far as age (mean = 67.4 years, and median = 66.4 years, where the youngest patient was 50.16 years, and the oldest 86.58 years, respectively), mass (mean = 63.15 kg and median = 62.5 kg) and gender distribution is concerned (males 56–40%) and females (83–60%).

The distribution of acetylator phenotypes in the two groups is shown in Table 1. The subjects were classified as fast acetylators if their phenotype suggested they were fast eliminators (FF = $k > 0.57\ h^{-1}$), intermediate eliminators (FS = $0.31\ h^{-1} < k < 0.57\ h^{-1}$; 1.2 h $< t_{\frac{1}{2}}\beta < 2.2$ h) or slow eliminators (SS = $k < 0.31\ h^{-1}$, $t_{\frac{1}{2}}\beta > 2.2$ h) of INH from their blood circulation.

From Table 1, it is evident that there are fewer fast and intermediate acetylators, and a significantly higher proportion of slow acetylators ($\chi^2 = 6.10$; $P = 0.047$) in the cases, compared to the controls.

Genotype analysis was concordant with phenotype in all instances. Thus, the result obtained and presented in Table 1 was confirmed by two independent methods. The detailed genotype analysis is shown in Table 2.

Phenotypes are defined by the following pharmacokinetic parameters: fast eliminators (FF) = $k > 0.57\ h^{-1}$, $t_{\frac{1}{2}}\beta < 1.2$ h; intermediate eliminators (FS) = $0.31\ h^{-1} < k < 0.57\ h^{-1}$; 1.2 h $< t_{\frac{1}{2}}\beta < 2.2$ h; slow eliminators (SS) = $k < 0.31\ h^{-1}$, $t_{\frac{1}{2}}\beta > 2.2$ h.

A comparison of the allele frequency (classified simply as F (fast) or S (slow)) between the cases and controls reveals a difference such that the slow allele confers a significant extra risk for cataract development (Table 3; $\chi^2 = 6$, 15, $P = 0.013$, OR = 1.57; CI 95% 1.09–2.27). An analysis of individuals classified as homozygous fast (FF) or heterozygous fast (FS) also reveals a similar effect ($P = 0.029$).

Table 1 Acetylator phenotype in subjects and controls

Phenotype	Cases (%)	Controls (%)
Fast	22 (15.8)	26 (22.8)
Intermediate	55 (39.6)	54 (47.4)
Slow	62 (44.6)	34 (29.8)
Totals	139	114

Table 2 Distribution of *N*-acetyl transferase phenotypes and genotypes in the case and control population

Patients ($n = 139$)	Phenotype* (genotype)	Alleles	Controls ($n = 114$)
22 (22)	Fast (FF)	NAT2*4/*4	26 (26)
36	Intermediate (FS)	NAT2*4/*5	20
16		NAT2*4/*6A	25
3		NAT2*4/*7	3
0 (55)		NAT2*4/*14A	6 (54)
19	Slow (SS)	NAT2*5	9
21		NAT2*5/*6A	10
4		NAT2*5/*7	4
2		NAT2*5/*14A	1
9		NAT2*6A/*6A	3
3		NAT2*6/*7	3
2		NAT2*6A/*14A	2
1		NAT2*7/*7	0
1		NAT2*7/*14A	1
0 (62)		NAT2*14A/*14A	1 (34)

F = fast allele (NAT2*4); S = generic for slow alleles (NAT2*5, NAT2*6, NAT2*6A, NAT2*7 and NAT2*14A); FF, FS and SS = homozygotic fast, heterozygotic intermediate and homozygotic slow phenotypes, respectively.

Discussion

The population in which the trial was conducted approaches ideal requirements since it is a culturally distinct, close-knit and stable ethnic entity. We believe that ethnic differences may be excluded as a confounding factor because the Coloured people of South Africa are recognized as a distinct ethnic group. The order of contribution to these people is likely to be the indigenous Khoi and San people, slaves brought from Malaysia, Indonesia and India in the 17th and 18th centuries, Europeans and black Africans. The genetic depth is approximately 10 generations [39,40]. A number of studies in our laboratories using this ethnic group as subjects have confirmed Hardy–Weinberg equilibrium and we have no reason to suspect any stratification between controls and patients in this study. In this population the alleles coding for fast and slow acetylator function are both well represented with frequencies of F = 0.465, and S = 0.535. These frequencies are reflected in comparable proportions of fast (FF) and slow (SS) acetylator individuals in the normal population: FF = 26%, and SS = 34%. The results show that individuals with 'slow' NAT2 alleles (particularly homozygotes) are clearly at greater risk for developing cataracts; however the reason, or reasons, for the association are at best speculative at this time, since acetylation is not generally known to occur in the human eye, whilst it has been reported to occur in bovine lens and in rabbit eye [41,42]. It is reasonable to assume that the correlations we observed are not merely co-incidental, and that progressive lens opacification is the result of additive damage caused by specific interactions at molecular lens protein level.

The most compelling possibility on available evidence is the role that acetylator function plays in the disposition of a wide variety of potentially toxic nitrogen

Table 3 NAT2 allele distribution between cataract cases and controls

Allele	Cases	Controls
F	99	106
S	179	122

containing molecular species, specifically those possessing primary amine or hydrazine moieties [43–46].

It is reasonable to assume that inefficient systemic disposition increases exposure of tissues, such as the lens, to the harmful effects of reactive molecules. Furthermore, it is possible that these molecules interact with vulnerable components in the lens, causing denaturation of proteinaceous molecules and, in time, lens opacification [47]. This possibility is supported by the observation that exposure to nitrogen containing chemicals, such as those occurring in tobacco smoke, has been associated with an increased risk of cataractogenesis [48]. Intuitively, slow acetylators would therefore run a greater risk of damage.

A second possibility resides in the ongoing post-translational modification of functional peptides in the lens, such as α-crystallin [48,49]. The latter is the most plentiful, and hence the single most important, component of the lens proper [50], particularly since the lens continues to grow throughout the life-span of the individual. α-crystallin has *N*-terminal methionine residues which are acetylated [51], and it is important to note that unlike proteins in most organs, there is no rapid turnover of α-crystallin in the eye [52]. Thus failure to acetylate α-crystallin in fetal or early life may predispose one to later pathologies. As such, acetylator function may have a protective role to play in the maintenance of lens transparency. However, a connection between systemic polymorphic NAT2 catalysed function, and local acetylation function in the lens, has not as yet been demonstrated. Causal interactions at molecular level remain to be elucidated.

Acknowledgements

The authors wish to thank Mr Paul Liebenberg for his invaluable research assistance in the collection of patient samples. Funding for this work came from NRF-THRIP, The Wellcome Trust, GlaxoSmithKline Action TB and the Harry Crossley Foundation.

References

1 Vatsis KP, Weber WW, Bell DA, Dupret JM, Evans DA, Grant DM, *et al.* Nomenclature for *N*-acetyltransferases. *Pharmacogenetics* 1995; **5**: 1–17.
2 Evans DAP. *N*-acetyltransferase. *Pharmacol Ther* 1989; **42**:157–234.
3 Lower GM Jr, Nilsson T, Nelson CE, Wolf H, Gamsky TE, Bryan GT. *N*-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ Health Perspect* 1979; **29**:71–79.
4 Cartwright RA, Glashan RW, Rogers HJ, Ahmad RA, Barham-Hall D, Higgins E, *et al.* Role of *N*-acetyltransferase phenotypes in bladder carcinogenesis: a pharmacogenetic epidemiological approach to bladder cancer. *Lancet* 1982; **2**:842–846.
5 Risch A, Wallace DM, Bathers S, Sim E. Slow *N*-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer. *Hum Mol Genet* 1995; **4**:231–236.
6 Brockmoller J, Cascorbi I, Kerb R, Roots I. Combined analysis of inherited polymorphisms in arylamine *N*-acetyltransferase 2, glutathione *S*-transferases M1 and T1, microsomal epoxide hydrolase and cytochrome P450 enzymes as modulators of bladder cancer risk. *Cancer Res* 1996; **56**:3915–3925.
7 Okkels H, Sigsgaard T, Wolf A and Autrup H. Arylamine *N*-acetyltransferase 1 (NAT1) and 2 (NAT2) polymorphisms in susceptibility to bladder cancer: the influence of smoking. *Cancer Epidemiol Biomark Prev* 1997; **6**:225–231.
8 Filiadis IF, Georgiou I, Alamanos Y, Kranas V, Giannakopoulos X, Lolis D. Genotypes of N-acetyltransferase 2 and risk of bladder cancer: a case-control study. *J Urol* 1999; **161**:1672–1675.
9 Borlak JT, Harsamy V, Schneble H, Haegele KD. PNAT and CYP2D6 gene polymorphism in epileptic patients. *Biochem Pharmacol* 1994; **48**:1717–1720.
10 Siegmind W, Fengler JD, Franke G, Zschiesche M, Eike O, Eike E, *et al.* *N*-acetylation and debrisoquine hydroxylation polymorphisms in patients with Gilbert's syndrome. *Br J Clin Pharmacol* 1991; **32**:467–472.
11 Baranova H, Canis M, Ivaschenko T, Albuisson E, Bothorishvili R, Bararov V, *et al.* Possible involvement of arylamine *N*-acetyltransferase 2, glutathione *S*-transferases M1 and T1 genes in the development of endometriosis. *Mol Hum Reprod* 1999; **5**:636–641.
12 Longuemaux S, Delomenie C, Gallou C, Mejean A, Vincent-Viry MN, Bouvier R, *et al.* Candidate genetic modifiers of individual susceptibility to renal cell carcinoma: a study of polymorphic human xenobiotic-metabolizing enzymes. *Cancer Res* 1999; **59**:2903–2908.
13 Hirvonen A, Pelin K, Tammilehto L, Karjalainen A, Mattson K, Linnainmaa K. Inherited GSTM 1 and NAT2 defects as concurrent risk modifiers in asbestos-related human malignant mesothelioma. *Cancer Res* 1995; **55**:2981–2983.
14 Hirvonen A, Saarikoski ST, Linnainmaa K, Loskinen K, Husgafvel-Puriainen K, Mattson K, *et al.* Glutathione *S*-transferase and *N*-acetyltransferase genotypes and asbestos-associated pulmonary disorders. *J Natl Cancer Inst* 1996; **88**:1853–1856.
15 Morita S, Yano M, Tsujinaka T, Ogawa A, Taniguchi M, Kaneko K, *et al.* Association between genetic polymorphisms of glutathione *S*-transferase P1 and *N*-acetyltransferase 2 and susceptibility to squamous cell carcinoma of the esophagus. *Int J Cancer* 1998; **79**:517–520.
16 Wikman H, Piirila P, Kaaria K, Rosenburg C, Norppa H, Hirvonen A. Role of NAT1 and NAT2 polymorphisms in individual susceptibility to isocyanate-induced asthma. *First International Workshop on the Arylamine N-transferases*, Queensland, Australia, 1998, Oct 22–24, (Abstract 12).
17 Zielinska E, Niewiarowski W, Bodalski J, Stanczyk A, Bolanowski W, Robowski G. Arylamine *N*-acetyltransferase gene mutations in children with allergic diseases. *Clin Pharmacol Ther* 1997; **62**:635–642.
18 von Schmiedeberg S, Fritsche E, Ronnau AC, Specker C, Golka K, Richter-Hintz D, *et al.* Polymorphisms of the xenobiotic-metabolizing enzymes CYP1A1 and NAT2 in systemic sclerosis and lupus erythematosus. *Adv Exp Med Biol* 1999; **455**:147–152.
19 Lang NP, Chu DZ, Hunter CF, Kendall DC, Flammang TJ, Kadlubar FF. Role of aromatic amine acetyltransferase in human colorectal cancer. *Arch Surg* 1986; **121**:1259–1261.
20 Ilett KF, David BM, Dethon P, Castleden WM, Kwa R. Acetylator phenotype in colorectal carcinoma. *Cancer Res* 1987; **47**:1466–1469.
21 Gil JP, Lechner MC. Increased frequency of wild-type arylamine-*N*-acetyltransferase allele NAT2*4 homozygotes in Portuguese patients with colorectal cancer. *Carcinogenesis* 1998; **19**:37–41.
22 el-Yazigi A, Johansen K, Raines DA, Dossing M. *N*-Acetylation polymorphism and diabetes mellitus among Saudi Arabians. *J Clin Pharmacol* 1992; **32**:905–910.
23 Cascorbi I, Brockmoller J, Mrozikiewicz PM, Bauer S, Loddenkemper R, Roots I. Homozygous rapid arylamine *N*-acetyltransferase genotype as a susceptibility factor for lung cancer. *Cancer Res.* 1996; **56**:3961–3966.
24 Pai P, Hindell P, Stevenson A, Mason H, Bell GM. Genetic variants of microsomal metabolism and susceptibility to hydrocarbon-associated glomerulonephritis. *Q J Med* 1997; **90**:693–698.
25 Philip PA, Rogers HJ, Millis RR, Rubens Rd, Cartwright RA. Acetylator status and its relationship to breast cancer and other diseases of the breast. *Eur J Cancer Clin Oncol* 1987; **23**:1701–1706.
26 Henning S, Cascorbi I, Munchow B, Jahnke V, Roots I. Association of arylamine *N*-acetyltransferases NAT1 and NAT2 genotypes to laryngeal cancer risk. *Pharmacogenetics* 1999; **9**:103–111.
27 Hadasova E, Brysova V, Kadlcakova E. *N*-Acetylation in healthy and diseased children. *Eur J Clin Pharmacol* 1990; **39**:43–47.
28 Meyer D, Liebenberg P. Cataract Etiology. In: Agarwal S, Apple DJ, Agarwal A, Buratto L, Alio JL, Pandey SK, Agarwal A (editors). *Textbook of Ophthalmology*. New Delhi: Jaypee Brothers, 2002; pp. 1587–1619.
29 Shambhu DV, Palaniyandi SD, Ali HA. Oxygen radicals in the pathogenesis of cataracts. In: *Nutritional and Environmental Influences on the Eye*. London: CRC Press, 1999.

30 Marais JS. *The Cape Coloured People: 1652 to 1932*. Johannesburg: Witwatersrand University Press, 1957.
31 Parkin D.P, Vandenplas M, Donald PR, Vandenplas S, Botha FJH, Seifart HI, *et al*. Distribution of acetylator genotypes in the coloured population of the Western Cape region of South Africa. *S Afr J Sci* 1998; **94**: 404–406.
32 Parkin DP, Vandenplas S, Botha FJH, Vandenplas M, Seifart HI, van Helden PD, *et al*. Trimodality of isoniazid elimination: Phenotype and genotype in patients with tuberculosis. *Am J Respir Crit Care Med* 1997; **155**:1717–1722.
33 Stryjek-Kaminska D, Malczewski B, Kopec A, Rowinska-Marcinska K. Acetylator phenotypes in diabetes mellitus. *Acta Diabetol Lett* 1988; **25**:41–48.
34 Meyer D, Maritz FJ, Liebenberg PH, Parkin DP, Burgess LJ. Cortical lens opacities in the young patient–an indication for a lipogram? *S Afr Med J* 2001; **91**:520–524.
35 Bialas MC, Routledge PA. Adverse effects of corticosteroids. *Adverse Drug React Toxicol Rev* 1998; **17**:227–235.
36 Ajamian PC. Traumatic cataract. *Optom Clin* 1993; **3**:49–56.
37 Seifart HI, Gent WL, Parkin DP, van Jaarsveld PP, Donald PR. High-performance liquid chromatographic determination of isoniazid, acetylisoniazid and hydrazine in biological fluids. *J Chromatogr B* 1995; **674**:269–275.
38 Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 1988; **16**:1215.
39 Nurse, GT, Weiner JS, Jenkins T. *The Peoples of Southern Africa and their Affinities*. Oxford: Clarendon Press, 1985.
40 Ncayiyana DJ. 'African' is a rainbow word. *S Afr Med J* 1997; **87**: 1 (Editorial).
41 Campbell DA, Schoenwald RD, Duffel MW, Berfknecht CF. Characterization of arylamine acetyltransferase in the rabbit eye. *Invest Ophthalmol Vis Sci* 1991; **32**:2190–2200.
42 Gaudet SJ, Razavi P, Caiafa GJ, Chader GJ. Identification of arylamine *N*-acetyltransferase activity in the bovine lens. *Curr Eye Res* 1995; **14**:873–877.
43 Holdiness MR. Clinical pharmacokinetics of the antituberculosis drugs. *Clin Pharmacokinet* 1984; **9**:511–544.
44 Clark DWJ. Genetically determined variability in acetylation and oxidation: therapeutic implications. *Drugs* 1985; **29**:342–375.
45 Weber WW. *The Acetylator Genes and Drug Response*. Oxford: Oxford University Press Inc., 1987.
46 Smith CAD, Smith G, Wolf CR. Genetic polymorphisms in xenobiotic metabolism. *Eur J Cancer* 1994; **30**:1921–1935.
47 West S, Munoz B, Schein OD, Vitale S, Maguire M, Taylor HR, Bressler NM, et al: Cigarette smoking and risk for progression of nuclear opacities. *Arch Ophthalmol* 1995; **113**:1377–1380.
48 Paterson CA, Delamere NA. The lens: In: Hart WA (editor). *Adler's Physiology of the Eye*, 9th edition. London: Mosby, 1992.
49 Bron AJ, Tripathi RC, Tripathi BJ. The lens and zonules. In: Bron AJ, Tripathi RC, Tripathi BJ (editors) *Wolff's Anatomy of the Eye and Orbit*, 8th edition. London: Chapman & Hall, 1997.
50 Hoenders HJ, Bloemendal H. Aging of lens proteins. In: Bloemendal H (editor), *Molecular and Cellular Biology of the Eye Lens*. New York: John Wiley, 1981.
51 Augusteyn RC. Protein modification in cataract: possible oxidative mechanisms. In: Duncan C (editor). *Mechanisms of Cataract Formation in the Human Lens*. London: Academic Press, 1981.
52 Groenen PJTA, Merck KB, De Jong WW, Bloemendal H. Structure and modifications of the junior chaperone α-crystallin. *Eur J Biochem* 1994; **225**:1–19.

CHAPTER 9

Isoniazid Pharmacokinetics in children treated for Respiratory Tuberculosis

Isoniazid pharmacokinetics in children treated for respiratory tuberculosis

H.S. Schaaf[1‡], D.P. Parkin[2], H.I. Seifart[2], C.J. Werely[3], P.B. Hesseling[1], P.D. van Helden[3], J.S. Maritz[4], and P.R. Donald[1]

Departments of [1]Paediatrics and Child Health, [2]Pharmacology, and [3]Medical Biochemistry, Medical Faculty, Stellenbosch University, P.O. Box 19063 Tygerberg 7505, South Africa.

[4]Biostatistical Unit of the South African Medical Research Council, Francie van Zijl Drive, Tygerberg, South Africa.

[‡]Author for correspondence (email: hss@sun.ac.za)

Archives of Diseases in Childhood (2005) **90**(6)**:**614-618.

Précis

Isoniazid (INH) is currently the only recommended prophylactic treatment for children with latent TB. However, there is a lack of data on the pharmacokinetics of INH in children, and it has been suggested that adhering to the dosing guidelines recommended for adults (based on body weight) may not be suitable, due to the larger liver-to-body mass ratio in children. We therefore investigated the pharmacokinetics of INH in children with respect to their *NAT2* genotype.

Our results showed that the metabolism of INH in children follows the expected trimodal pattern, according to the fast (FF), Intermediate (FS), and slow (SS) genotype-classes. Secondly, the rate of INH metabolism in children was also faster compared to adults. Furthermore, the highest rate of INH elimination was observed in children less than 5 years of age, with older children having a slower rate of elimination by comparison. More importantly, the metabolism of INH was not related to the weight of the child. This study investigated the pharmacology of INH elimination in children based upon a once-off dose of INH (of 10 mg/kg body weight), which proved to be sub-optimal in the FF acetylators, who achieved a peak serum concentration of less than 2 mg/l. These observations support the hypothesis that the rate of INH metabolism may be ascribed to the higher (total) metabolic activity of the relatively large liver in children compared to adults. Hence the proposed optimal dosing for children should be based on a calculation of the surface area of the body, rather than the mass. These results also provide support for the recommendation that generally children should receive a higher dose of INH than adults.

It would therefore be interesting to test a range of INH doses in children (of known acetylator status) to determine the INH dose for optimum efficacy, as well as validate these results in children receiving co-administered antibiotics, as recommended by the DOTs regimen for TB treatment.

Sample Source:- Adolescent samples were sourced from the Tygerberg Hospital, Western Cape Province. Sample cohort used in this study only.

All NAT2 genotyping was performed by CJW, who wrote the experimental methodology in the manuscript, and contributed to the writing of the manuscript.

ORIGINAL ARTICLE

Isoniazid pharmacokinetics in children treated for respiratory tuberculosis

H S Schaaf, D P Parkin, H I Seifart, C J Werely, P B Hesseling, P D van Helden, J S Maritz, P R Donald

Arch Dis Child 2005;**90**:614–618. doi: 10.1136/adc.2004.052175

See end of article for authors' affiliations

Correspondence to: Dr H S Schaaf, Paediatrics and Child Health, Stellenbosch University, PO Box 19063, 7505 Tygerberg, South Africa; hss@sun.ac.za

Accepted 9 August 2004

Aims: To define the pharmacokinetics of isoniazid (INH) in children with tuberculosis in relation to the N-acetyltransferase 2 (NAT2) genotype.

Methods: The first order elimination rate constant (*k*) and area under the concentration curve (AUC) were calculated in 64 children <13 years of age (median 3.8) with respiratory tuberculosis from INH concentrations determined 2–5 hours after a 10 mg/kg INH dose. The NAT2 genotype was determined; 25 children were classified as homozygous slow (SS), 24 as heterozygous fast (FS), and 15 as homozygous fast (FF) acetylators.

Results: The mean (SD) *k* values of the genotypes differed significantly from one another: SS 0.254 (0.046), FS 0.513 (0.074), FF 0.653 (0.117). Within each genotype a median regression of *k* on age showed a significant decrease in *k* with age. The mean (SD) INH concentrations (mg/l) two hours after INH administration were SS 8.599 (1.974), FS 5.131 (1.864), and FF 3.938 (1.754). A within genotype regression of 2-hour INH concentrations on age showed a significant increase with age. A within genotype regression of 3-hour, 4-hour, and 5-hour concentrations on age also showed a significant increase with age in each instance. In ethnically similar adults, mean (SD) 2-hour INH concentrations (mg/l) for each genotype were significantly higher than the children's: SS 10.942 (1.740), FS 8.702 (1.841), and FF 6.031 (1.431).

Conclusions: Younger children eliminate INH faster than older children and, as a group, faster than adults, and require a higher mg/kg body weight INH dose to achieve serum concentrations comparable to adults.

More than 50 years after its introduction as an antituberculosis agent,[1] isoniazid (INH) continues to form the cornerstone of all "first line" antituberculosis regimens and remains the only agent recommended for tuberculosis chemoprophylaxis in children.[2 3] INH is most valued for its powerful bactericidal effect against the metabolically active organisms most commonly encountered in the sputum of adults with cavitating pulmonary tuberculosis[4] and is the most valuable agent for preventing the development of resistance in companion agents.[5] There is also evidence that INH suppresses the growth of non-multiplying organisms and that prolonged exposure of such organisms to concentrations above the minimal inhibitory concentration (MIC) leads to bacteriolysis.[6 7]

INH is well absorbed from the gastrointestinal tract, but is subject to significant first pass metabolism that may impact on its systemic concentrations. INH does not bind appreciably to plasma proteins, crosses membranes readily, and distributes into a compartment that approximates the total body water.[8] Very little of the parent compound is excreted unchanged in the urine and the greater proportion is acetylated in the liver and the small intestine to acetylisoniazid prior to excretion in the urine. Acetylation capacity in any individual is genetically determined, and although the initial studies of INH pharmacokinetics seemed to indicate that INH was bimodally eliminated,[9] there was the suspicion that its elimination was in fact trimodal.[10] Modern molecular biology techniques and improved analytical methods have now established, beyond doubt, that INH is eliminated in accordance with a trimodal distribution of subtypes, fast (FF), intermediate (FS), and slow (SS), the fast (F) and slow (S) alleles being co-dominant.[11]

The pharmacokinetic characteristics of INH have been extensively studied in adults,[8] but data in respect of children, and especially younger children are limited. Where such data are available cognisance has not been taken of the genotype or of the trimodality of INH elimination.[12–14]

This study was undertaken to improve our understanding of INH pharmacokinetics in children being treated for tuberculosis, making use of improved analytical technology and advances in our understanding of the polymorphisms governing INH metabolism.

PATIENTS AND METHODS

The study was undertaken in the Western Cape Province of South Africa, an area with a particularly high incidence of tuberculosis (>600/100 000 population at the time of this study). Children less than 13 years of age with primary respiratory tuberculosis and some with abdominal tuberculosis were included in the study following informed written consent of the parent or legal guardian. If not already hospitalised for other reasons the children enrolled in the study were temporarily admitted to hospital on the morning of the study.

The age and weight of each child were recorded and the extent of pulmonary involvement and the presence of extrapulmonary tuberculosis noted. Children who were severely ill were excluded from the study. With appropriate counselling and written informed consent the human immunodeficiency virus (HIV) status of the children was assessed.

The INH used for the study was standard pharmaceutical grade in powder form obtained from Fluka Chemie AG (Buchs, Switzerland). The INH was accurately weighed to

Abbreviations: AUC, area under the concentration curve; FF, homozygous fast acetylator; FS, heterozygous fast acetylator; INH, isoniazid; NAT2, N-acetyltransferase 2; SS, homozygous slow acetylator

Table 1 Clinical features and special investigation results of the study children

Clinical feature or special investigation	Number n = 64	%
Age group		
0–2 years	18	28
>2–5 years	24	38
>5–13 years	22	34
Weight loss	50	78
Weight <3rd centile	31	48
Cough >2 weeks	39	61
Household TB contact	35	55
Mantoux test		
0–4 mm	13 (7 = HIV infected)	20
5–14 mm	6	9
⩾15 mm	41 (6 = HIV infected)	64
Not done	4	6
HIV infected	13	20
Chest radiograph		
Lymphadenopathy	44	69
Collapse/opacification	33	52
Pleural effusion	14	22
Miliary	8	13
Extra-pulmonary TB		
Peripheral lymph nodes	14	22
Pleural effusion	14	22
Miliary	8	13
TB meningitis stage 1	4	6
Pericardial effusion	1	2
Culture or histology confirmed tuberculosis	41	64

give a dose of 10 mg/kg according to the child's weight the previous day. The INH powder was dissolved in 5–10 ml of water and administered by one of the study personnel orally with a syringe or, in the case of very young children, through a nasogastric tube and washed down with water. A light breakfast was permitted 60–90 minutes later.

Four blood specimens of 1–1.5 ml each, taken at 2, 3, 4, and 5 hours post-dose, were collected in ethylenediaminetetraacetate (EDTA) coated tubes, chilled, and delivered on ice to the laboratory within one hour of taking the last specimen. INH concentrations were determined by the high performance liquid chromatography (HPLC) method of Seifart and colleagues.[15] A further single 3 ml sample was collected into an EDTA coated tube for DNA analysis.

NAT2 genotyping

Genomic DNA (gDNA) was prepared via the salting out procedure of Miller and colleagues.[16] This gDNA was subsequently analysed for the *NAT2*5*, *2*6*, *2*7*, *2*12*, *2*13*, and *2*14* alleles,[17] via a polymerase chain reaction (PCR) based strategy, as previously described.[11] Separate PCR aliquots were restricted with the *Msp*I, *Fok*I, *Kpn*I, *Taq*I, *Dde*I, and *Bam*HI restriction enzymes (according to the manufacturer's recommendations) to delineate the polymorphisms at nucleotide positions 191, 282, 481, 590, 803, and 857, respectively. According to the Vatsis nomenclature the wild type fast allele (F) is assigned as *NAT2*4*, *2*12*, or *2*13*.[17] These alleles confer normal enzyme activity on the NAT2 protein, while the mutant slow alleles (S), classified as *NAT2*5*, *2*6*, *2*7*, and *2*14* in humans, confer a decreased enzyme activity on the NAT2 protein.

The *T341C* mutation of the *NAT2*5* allele was typed via an allele specific PCR protocol, employing the confronting primer PCR method of Hamajima and colleagues.[18] In this case the standard PCR mixture contained two primer sets, primer set I [5′-^{876nt}TTAGAGGCTATTTTTGAT CACA^{897nt}-3′ and 5′-^{1081nt}ATGTAATTCCTGCCGTCAG^{1063nt}-3′], which initiates amplification in the case of the 341C allele (a 187bp product), and primer set II [5′-^{1045nt}TTCTCCTGCAGGTGAC CAT^{1063NT}-3′ and 5′-^{1368nt}AAGATGTTGGAGA CGTCT GC^{1349nt}-3′], which only amplifies in the case of the 341T allelic sequence (a 323bp product). In addition, outermost primers [^{876nt}TTAGAG GCTATTTTTGATCACA^{897nt} and ^{1368nt}AAGATGTTGGAGACGTCTGC^{1349nt}] also amplify a gene specific PCR product (475bp) in the reaction mixture which serves as an internal amplification control.

The *Kpn*I, *Bam*HI, and *Msp*I generated profiles were subsequently also cleaved with *Pst*I, to facilitate the analyses of these DNA profiles. This second cleavage generated a profile of bands less than 500 bp in size which enabled their resolution via 5% polyacrylamide gel electrophoresis (PAGE), using the Mini-Protean II gel system (Bio-Rad Laboratories, USA); DNA bands were visualised by silver staining, as described by Bassam and colleagues.[19] The *Pst*I restriction sites occur at nucleotide positions 1050 and 1347 of the *NAT2* gene sequence (GenBank accession number D10870); we have found these sites to be conserved in all samples that we have analysed thus far (more than 700). The use of this second enzymatic restriction improves the sensitivity of analysis of the polymorphisms at 481, 857, and 191 respectively.

Pharmacokinetic parameters

The apparent first order elimination rate constants (k, h^{-1}) of the individual patients were calculated from the linear regression of $\ln(C_t)$ and t, where $\ln(C_t)$ is the natural logarithm of the concentration at time t; observations of C_t were made at 2, 3, 4, and 5 hours after oral administration of 10 mg/kg body weight of INH. The area under the curve (AUC) over the interval 2–5 hours after the dose was calculated by integration of the AUC of each of the constituent subintervals (2–3, 3–4, and 4–5 hours after the dose) by standard methods.[20]

Statistical methods

Tests of homogeneity of group means were performed by one way analysis of variance. The general linear model approach was used for data subject to more than one classification criterion. The association between quantitative variables was examined by regression; in the case of the relation between age and k values the heteroscedasticity of the data required

Table 2 The mean first order elimination rate constant (k), the area under the curve (AUC) during the period 2–5 hours after dosing, and the mean serum INH concentrations at 2, 3, 4, and 5 hours after dosing with 10 mg/kg isoniazid (INH)

			Mean INH concentration (SD) (mg/l)			
Genotype (n)	K (SD) (h^{-1})	AUC (SD) (mg/l/h)	2 h after dosing	3 h after dosing	4 h after dosing	5 h after dosing
SS (25)	0.254 (0.046)	18.356 (4.692)	8.599 (1.974)	6.585 (1.615)	5.099 (1.355)	4.007 (1.179)
FS (24)	0.513 (0.074)	8.246 (3.349)	5.131 (1.864)	3.168 (1.294)	1.955 (0.885)	1.178 (0.565)
FF (15)	0.653 (0.117)	5.371 (3.076)	3.938 (1.754)	2.045 (1.060)	1.125 (0.646)]	0.642 (0.426)
F(df,df)	88.92 (2.61)	58.42 (2.55)	33.39 (2.59)	61.94 (2.61)	84.91 (2.61)	0.642 (2.57)
p value	<0.0005	<0.0005	<0.0005	<0.0005	<0.0005	<0.0005

SD, standard deviation; SS, homozygous slow acetylators; FS, heterozygous fast acetylators; FF, homozygous fast acetylators.

Table 3 Significance of a straight line regression of INH concentrations on age with a common slope for genotype

Time after INH dosing (hours)	Ordinates at mean age SS	FS	FF	Slope	p*
2	8.622	5.078	3.882	0.278	<0.0005
3	6.650	3.128	2.001	0.214	<0.0005
4	5.150	1.924	1.092	0.166	<0.0005
5	4.048	1.129	0.604	0.136	<0.0005

SS, homozygous slow acetylators; FS, heterozygous fast acetylators; FF, homozygous fast acetylators.
*Significance of the slope coefficient.

the use of median regression. This is essentially robust regression that entails the fitting of parallel straight lines to the three groups of data.[21]

The Institutional Review Board of the Faculty of Health Sciences, Stellenbosch University, approved the study. The parents or legal guardians of all of the children gave written informed consent for their children's participation in the study.

RESULTS

Sixty four children with a median age of 3.8 years (lower quartile 1.8 years, upper quartile 7.8 years) were included in the study. Table 1 summarises details of age and radiological and clinical features. Although 50 (78%) children had experienced recent weight loss, confirmed by reference to a "Road to Health" card, only 31 (48%) had a weight for age of less than 3rd centile (National Center for Health Statistics, USA). Thirteen (20%) children were HIV infected and 15 (23%) children suffered from abdominal tuberculosis. In 41 children (64%) tuberculosis was confirmed by isolation of *M tuberculosis* from one or more clinical specimens or identification of acid fast bacilli following tissue biopsy.

Twenty five children (39%) with median age of 2.8 years were genotyped as homozygous slow acetylators of INH, 24 (38%) with median age of 3.9 years heterozygous fast acetylators, and 15 (23%) with median age of 4.1 years as homozygous fast acetylators. With respect to the variable age, the differences between the genotype groups are not statistically significant.

Table 2 presents the mean first order elimination rate constant (k), the mean area under the concentration versus time curve (AUC) for the period 2–5 hours after dosing, and the mean INH concentration at 2, 3, 4, and 5 hours after administration of INH. The means of each of the variables differed significantly between genotypes ($p < 0.0005$ in each instance). In the case of the variable k the group variance increases with group means and for formal testing of homogeneity of group means the k values were transformed to logarithms; this had the effect of stabilising the variances. The ratio of the largest to the smallest variance gave observed $F(25.22) = 1.49$; this conservative test gives $p = 0.18$. A formal test of normality of residuals was not performed but dotplots indicated no obvious departure from normality. Although one way ANOVA is sensitive to lack of homoscedasticity it is robust against departures from normality. A Kruskal-Wallis test applied to k (or its logarithms) results in a similar χ^2 (49.17; degrees of freedom = 2, $p < 0.0005$). Similar checks of the results relating to other variables in table 2, confirmed the statistical significance statements. Confidence limits (95%) for the location of the log transformed data are: SS (0.234 to 0.268), FS (0.477 to 0.538), FF (5.93 to 0.700). They are clearly well separated.

At the extremes the two hour INH concentrations varied from a high of 12 mg/l in a homozygous slow acetylator to a low of <2 mg/l in a homozygous fast acetylator. Within each genotype the means of k, AUC, and 2-hour, 3-hour, 4-hour, and 5-hour INH concentrations were not significantly associated with either HIV infection or the presence of abdominal tuberculosis.

With regard to the relation between age and k a regression of median values, taking into account heteroscedasticity of the data, confirms a significant decline in k with increasing age for each genotype. The results of fitting straight line regressions, k on age, with common slope parameter, were: slope = −0.00521, standard error 0.00114, giving $p < 0.001$.[21] The fitted median values at the common mean age are 0.248, 0.500, 0.620 respectively for the three genotypes; they differ significantly from each other according to a global test ($p < 0.0005$), and the two values closest to each other, FS and FF also differ significantly.[21] Table 3 summarises the result of fitting straight line regressions, INH concentration at each of the time points, 2-hours, 3-hours, 4-hours, and 5-hours on age with a common slope for the genotypes, and confirms a significant rise in each of the 2-hour, 3-hour, 4-hour, and 5-hour INH concentrations with age ($p < 0.0005$ in each instance).

In order to explore the relation of INH elimination and age further we compared the findings in this group of children with those obtained by our research group in a similar ethnic population of adults who were also genotyped and phenotyped after receiving 5 mg/kg and 10 mg/kg doses of INH.[11] Among the adult patients INH concentrations were

Table 4 Adult patients (n = 60): the mean first order elimination rate constant (k), the area under the curve (AUC) during the period 2–5 hours, the mean serum concentrations at 2 and 3 hours after dosing with 10 mg/kg, and the calculated INH concentrations 4 and 5 hours after dosing

Genotype	k (SD) (h^{-1})	AUC (SD) (mg/l/h)	Mean INH concentration (SD) (mg/l) 2 h after dosing	3 h after dosing	4 h after dosing	5 h after dosing
SS	0.193 (0.026)	24.870 (4.077)	10.942 (1.740)	8.943 (1.448)	7.433 (1.268)	6.169 (1.141)
FS	0.430 (0.080)	15.338 (4.018)	8.702 (1.841)	5.789 (1.446)	3.991 (1.225)	2.631 (1.015)
FF	0.678 (0.056)	8.139 (2.167)	6.031 (1.431)	3.043 (0.896)	1.748 (0.510)	0.877 (0.257)

SD, standard deviation; SS, homozygous slow acetylators; FS, heterozygous fast acetylators; FF, homozygous fast acetylators.

Table 5 Results of the ANOVA of variable *k*

Factor	df	F	p
Genotype	2, 118	263.09	<0.0005
Adult/child	1, 118	6.79	0.010
Gen*adult/child	2, 118	3.96	0.022

Table 6 Results of the ANOVA of variable A

Factor	df	F	p
Genotype	2, 118	62.68	<0.0005
Adult/child	1, 118	62.51	<0.0005
Gen*adult/child	2, 118	1.68	0.191

determined at 0.5, 1, 2, 3, 4.5, and 6 hours after dosing. Table 4 presents the adults' mean INH concentrations at 2 and 3 hours after dosing, together with the calculated 4- and 5-hour INH concentrations, *k* values, and AUC for the period 2–5 hours after dosing. These means may be compared with the paediatric data in table 2. However, for a more formal comparison of the two data sets we recall the basic model relating concentration and time of observation according to which log (concentration) is a linear function of time. The slope coefficient is the rate constant k, and the implication is that the data of any individual can be summarised in this constant and the intercept of the linear function. Equivalently the data can be summarised in *k* and the fitted ordinate at any chosen time. For this analysis we use the ordinate at 2 hours, and for convenience refer to it as A.

Genotype and adult/child status are treated as factors with three and two levels, respectively, in a two way analysis of variance to test for the significance of these factors. Table 5 shows the results of the ANOVA of variable *k*.

All effects are clearly highly significant. As a check a rank regression analysis was performed, as implemented in MINITAB, and essentially the same results were obtained. Inspection of tables 2 and 4 shows that the significant interaction is explained by the mean values of *k* being different between children and adults at genotypes SS and FS, but not at FF. To confirm, the Bonferroni confidence intervals with global confidence coefficient 0.95 for the differences between the adult and child means are: SS (0.0029 to 0.1201), FS (0.0274 to 0.1386), FF (−0.1017 to 0.0517). According to the ANOVA the overall means of the genotypes differ significantly. Bonferroni intervals, as above, for the differences are, SS–FS (−0.2826 to −0.2020), SS–FF (−0.4855 to −0.3895), FS–FF (−0.2423 to −0.1481).

Table 6 present the results of an analysis of variance of variable A. The genotype and adult/child main effects are clearly significant. These results have also been confirmed by a rank regression analysis. The Bonferroni confidence intervals with global confidence coefficient 0.95 for the genotype mean differences are: SS–FS (1.628 to 3.402), SS–FF (3.668 to 5.782), FS–FF (1.172 to 3.248). The conclusion is that average levels of concentration differ between adults and children, and also between genotypes.

Finally, as further confirmation of the significance of the genotype and adult/child differences a multivariate ANOVA was performed with the joint response variables *k* and A. For the test of the main effect of factor adult/child the Wilk F-statistic is $F(2,117) = 31.10$, $p < 0.0005$; for the test of the main effect of factor genotype the Wilk F-statistic is $F(4,234) = 80.28$, $p < 0.0005$; and for the interaction it is $F(4,234) = 2.40$, $p = 0.051$.

After transforming the children's weight for age to *z* scores, analysis of variance did not reveal any significant association of the pharmacokinetic parameters with body weight.

DISCUSSION

The results of this study illustrate once again the considerable differences in exposure to INH that exist between homozygous slow acetylators of INH and the heterozygous and homozygous fast acetylators, and show that these differences also exist in children. Our data also confirm that younger children eliminate INH faster than older children and in a trimodal model of INH elimination there is a significant age related decline in the first order elimination rate constant (k, h^{-1}) with age in all three genotypes. Furthermore the exposure of the children to INH, as reflected by the first order elimination rate constant, AUC for the period 2–5 hours after dosing, and INH concentrations at different time intervals after dosing, is significantly less than that of a group of adults drawn from the same population and receiving the same mg/kg body weight dose of INH. These findings, taking into account the NAT2 genotype for the first time, confirm the suggestions of earlier workers, based on phenotypification, that younger children eliminate INH faster than older children, and children, as a group, faster than adults.[22–25] The significantly faster elimination of INH by infants and younger children has been ascribed to the relatively greater mass of the liver in proportion to total body weight and it has been proposed that more optimal doses would be calculated on the basis of body surface area rather than body weight. Given that INH will most often be used in developing countries under programme conditions it is unlikely that this will be possible.

The normal range of INH concentrations two hours after dosing has been given as 3–5 mg/l;[26] alternatively it has been suggested that a 3-hour concentration of 1.5 mg/l is desirable.[27] It is therefore noteworthy that seven (35%) of the homozygous fast acetylators had a 2-hour INH concentration of less than 3 mg/l and nine (45%) did not reach a 3-hour post-dose concentration of 1.5 mg/l. It is inevitable that using a lower dose of 5 mg/kg body weight will lead to an even greater proportion of homozygous fast acetylators and probably a significant proportion of heterozygote fast acetylators failing to achieve the recommended concentrations.

The above findings provide justification for the official recommendations of some professional bodies that children should receive higher mg/kg doses of INH than adults. Thus while the International Union Against Tuberculosis and Lung Disease,[28 29] the World Health Organisation,[30] and the Joint Tuberculosis Committee of the British Thoracic Society[31] recommend an INH dosage of 4–6 mg/kg body weight, the American Academy of Pediatrics recommends an INH dose of 10–15 mg/kg body weight.[2] In debating the most appropriate dose of INH for use in children, some caution is necessary. Many of the studies documenting a satisfactory clinical response to an INH dose of 5 mg/kg emanate from populations with a predominance of homozygous slow acetylators; the relative therapeutic disadvantage of the heterozygote fast acetylator will also be concealed by the synergistic potential of a multidrug regimen. Under adverse circumstances, however, such as failure of full compliance, compromised absorption of INH itself, or its companion drugs, in particular rifampicin,[32] or the sequestration of organisms in poorly perfused lesions,[33–35] this relative disadvantage may be exposed. In the light of our findings, and in agreement with the recommendations of the American Academy of Pediatrics, we suggest that young children less than 5 years of age should receive an INH dose of at least 10 mg/kg to ensure that the faster acetylators of INH are exposed to adequate serum concentrations of INH.

Authors' affiliations
H S Schaaf, P B Hesseling, P R Donald, Department of Paediatrics and Child Health, Faculty of Health Sciences, Stellenbosch University and Tygerberg Children's Hospital, South Africa
D P Parkin, H I Seifart, Department of Pharmacology, Faculty of Health Sciences, Stellenbosch University, South Africa
C J Werely, P D van Helden, Department of Medical Biochemistry, Faculty of Health Sciences, Stellenbosch University, South Africa
J S Maritz, Biostatistics Unit of the South African Medical Research Council, South Africa

Financial assistance: Harry and Doris Crossley Foundation (HS Schaaf) and National Research Foundation (PD van Helden)

Competing interests: none declared

REFERENCES

1 **Fox HH**. The chemical approach to the control of tuberculosis. *Science* 1952;**116**:129–34.
2 **American Academy of Pediatrics**. Tuberculosis. In: Pickering LK, ed. *Red Book: 2003 Report of the Committee on Infectious Diseases*, 26th edn. Elk Grove Village, IL: American Academy of Pediatrics, 2003:642–60.
3 **Centers for Disease Control and Prevention**. Targeted tuberculin testing and treatment of latent tuberculosis infection. *MMWR* 2000;**49**:1–51.
4 **Donald PR**, Sirgel FA, Botha FJ, *et al*. The early bactericidal activity of isoniazid related to its dose size in pulmonary tuberculosis. *Am J Respir Crit Care Med* 1997;**156**:895–900.
5 **Mitchison DA**. Role of individual drugs in the chemotherapy of tuberculosis. *Int J Tuberc Lung Dis* 2000;**4**:796–806.
6 **Crowle AJ**, Sbarbaro JA, May MH. Effects of isoniazid and ceforamide against virulent tubercle bacilli in cultured human macrophages. *Tubercle* 1988;**69**:15–25.
7 **McCune RM**, Feldman FM, Lambert HP, *et al*. Microbial persistence 1. The capacity of tubercle bacilli to survive in mouse tissues. *J Exp Med* 1966;**123**:445–68.
8 **Weber WW**, Hein HW. Clinical pharmacokinetics of isoniazid. *Clin Pharmacokinet* 1979;**4**:401–22.
9 **Evans DA**, Manley KA, McKusick VA. Genetic control of isoniazid metabolism in man. *BMJ* 1960;**2**:485–91.
10 **Sunahara S**, Urano M, Agawa M. Genetic and geographical studies on isoniazid inactivation. *Science* 1961;**134**:1530–1.
11 **Parkin DP**, Vandenplas S, Botha FJH, *et al*. Trimodality of isoniazid elimination. *Am J Respir Crit Care Med* 1997;**155**:1717–22.
12 **Pariente-Khayat A**, Rey E, Gendrel D, *et al*. Isoniazid acetylation metabolic ratio during maturation in children. *Clin Pharmacol Ther* 1997;**62**:377–83.
13 **Rey E**, Gendrel D, Treyluyer JM, *et al*. Isoniazid pharmacokinetics in children according to acetylator phenotype. *Fund Clin Pharmacol* 2001;**15**:355–9.
14 **Seth V**, Beotra A. Antituberculosis drugs. II: Clinical pharmacokinetics in Indian children. In: Seth V, eds. *Essentials of tuberculosis in children*. Delhi, India: Jaypee Brothers Medical Publishers, 1997:291–303.
15 **Seifart HI**, Gent WL, Parkin DP, *et al*. High-performance liquid chromatographic determination of isoniazid, acetylisoniazid and hydrazine in biological fluids. *J Chromatograph B* 1995;**674**:269–75.
16 **Miller SA**, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
17 **Vatsis KP**, Weber WW, Bell DA, *et al*. Nomenclature for N-acetyltransferases. *Pharmacogenetics* 1995;**5**:1–17.
18 **Hamajima NT**, Saito T, Matsuo K, *et al*. Polymerase chain reaction with confronting two-pair primers for polymorphism genotyping. *Jpn J Cancer Res* 2000;**91**:865–8.
19 **Bassam BJ**, Caetano-Anollés G, Gresshoff PM. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem* 1991;**196**:80–3.
20 **Weber WW**. Determination of the human acetylator status. In: *The acetylator genes and drug response*. New York: Oxford University Press, 1987:151–73.
21 **Maritz JS**, Donald PR, Schaaf HS, *et al*. The association of age and the elimination of isoniazid: an example of median regression with heteroscedasticity. *S Afr Statist J* 2003;**37**:191–202.
22 **Houin G**, Tillement JP. Consequences thérapeutiques de la mesure de l'indice d'inactivation de l'isoniazide au cours du traitement antituberculeux. *Therapie* 1980;**35**:597–605.
23 **Advenier C**, Saint-Aubin A, Scheinmann P, *et al*. Pharmacocinétique de l'isoniazide chez l'enfant. *Rev Fr Mal Respir* 1981;**9**:365–74.
24 **Bouveret JP**, Hanoteau J, Gerbeaux J, *et al*. Variations de l'indice d'inactivation de l'isoniazide au cours des traitements antituberculeux chez l'enfant. *Archiv Fr Pediatr* 1983;**40**:615–19.
25 **Lupasco I**, Algeorge G, Ghiu-Cimpeanu V. Taux d'isoniazide actif dans le serum des enfants tuberculeux. *Pneumon Coeur* 1965;**21**:1105–14.
26 **Peloquin CA**. Therapeutic drug monitoring: principles and application in mycobacterial infections. *Drug Ther* 1992;**22**:31–6.
27 **Vivien JN**, Thibier R, Grosset J, *et al*. Resultats précoces de l'isoniazido-thérapie en fonction du taux d'isoniazide actif dans le sérum. *Rev Tuberc* 1958;**22**:208–22.
28 **International Union Against Tuberculosis and Lung Disease**. Tuberculosis in children: guidelines for diagnosis, prevention and treatment. *Bull Int Union Tuberc Lung Dis* 1991;**66**:61–7.
29 **Enarson DA**, Rieder HL, Arnodottir T, *et al*. *Management of tuberculosis: a guide for low income countries*, 5th edn. Paris: IUATLD, 2000.
30 **World Health Organisation**. Standardized treatment regimen. In: *Treatment of tuberculosis: guidelines for national programmes*. WHO/CDS/TB2003.313. Geneva: WHO, 2003:27–38.
31 **Joint Tuberculosis Committee of the British Thoracic Society**. Chemotherapy and management of tuberculosis in the United Kingdom: recommendations 1988. *Thorax* 1998;**53**:536–48.
32 **Sahai J**, Gallicano K, Swick L, *et al*. Reduced plasma concentrations of antituberculosis drugs in patients with HIV infection. *Ann Intern Med* 1997;**127**:289–93.
33 **Iseman MD**, Madsen LA. Chronic tuberculous empyema with bronchopleural fistula resulting in treatment failure and progressive drug resistance. *Chest* 1991;**100**:124–7.
34 **Elliott AM**, Berning SE, Iseman MD, *et al*. Failure of drug penetration and acquisition of drug resistance in chronic tuberculous empyema. *Tubercle Lung Dis* 1995;**76**:463–7.
35 **Schoeman JF**, Morkel A, Seifart HI, *et al*. Massive posterior fossa tuberculous abscess developing in a young child treated for miliary tuberculosis. Possible role of very rapid acetylation of isoniazid. *Pediatr Neurosurg* 1998;**29**:64–8.

CHAPTER 10

Isoniazid plasma concentrations in children: Implications for international paediatric dosing guidelines

Isoniazid plasma concentrations in a cohort of South African children with tuberculosis: Implications for international paediatric dosing guidelines

Helen McIlleron[1], Marianne Willemse[4,5], Cedric J. Werely[6,7], Gregory D. Hussey[2,3], H. Simon Schaaf[4,5], Peter J. Smith[1] and Peter R. Donald[4,5]

[1]Division of Clinical Pharmacology, Department of Medicine, [2]Institute of Infectious Diseases and Molecular Medicine, and [3]School of Child and Adolescent Health, University of Cape Town, and [4]Tygerberg Children's Hospital, Cape Town, and [5]Department of Paediatrics and Child Health, [6]Division of Molecular Biology and Human Genetics, and [7]Medical Research Council Centre for Molecular and Cellular Biology, Faculty of Health Sciences, University of Stellenbosch, Stellenbosch, South Africa.

Clinical Infectious Diseases (2009) **48**(11):1547-1553

Précis

The elimination of isoniazid has previously been shown to be faster in children compared to adults, and yet the dosage guidelines for children have not been widely established. Here we investigated the effect of a range of INH doses in children, to determine the optimum dose range that would yield efficacies similar to that found in adults.

Pre-adolescent children, being treated for TB were enrolled in the study, and the *NAT2* genotypes in a group of 48 children were observed as 14% fast, 46% intermediate and 40% slow acetylators. Thirty-four children in the group were given a dose of 4-6 mg/kg of INH, whilst the remaining 14 received 8-12 mg/kg INH. In the first group, 76% of children [7/14 slow, 13/14 intermediate and 6/6 fast acetylators) recorded a minimum peak concentration below 3 mg/l (In adults a minimum INH peak concentration of 3 mg/l correlated with good treatment efficacy). In contrast, all of the children in the second group achieved this INH concentration level, irrespective of their *NAT2* genotype.

These results indicate that children given the standard 4-6 mg/kg dose of INH are at greater risk of exhibiting a sub-optimal treatment efficacy. Interestingly, the risk of INH-related hepatitis was not significantly different between these groups, indicating that the use of the higher INH dosage in children was relatively safe.

Sample Source:- Adolescent samples were obtained from the Brooklyn Hospital for Chest Diseases, Western Cape Province. Sample cohort used in this study only.

All NAT2 genotyping was done by CJW, who also contributed to the interpretation of results and writing of the manuscript.

MAJOR ARTICLE

Isoniazid Plasma Concentrations in a Cohort of South African Children with Tuberculosis: Implications for International Pediatric Dosing Guidelines

Helen McIlleron,[1] Marianne Willemse,[4,5] Cedric J. Werely,[6,7] Gregory D. Hussey,[2,3] H. Simon Schaaf,[4,5] Peter J. Smith,[1] and Peter R. Donald[4,5]

[1]Division of Clinical Pharmacology, Department of Medicine, [2]Institute of Infectious Diseases and Molecular Medicine, and [3]School of Child and Adolescent Health, University of Cape Town, and [4]Tygerberg Children's Hospital, Cape Town, and [5]Department of Paediatrics and Child Health, [6]Division of Molecular Biology and Human Genetics, and [7]Medical Research Council Centre for Molecular and Cellular Biology, Faculty of Health Sciences, University of Stellenbosch, Stellenbosch, South Africa

Background. In most countries with a high burden of tuberculosis, children with tuberculosis are prescribed isoniazid at dosages of 4–6 mg/kg/day, as recommended by international authorities.

Methods. We studied isoniazid concentrations in 56 hospitalized children (median age, 3.22 years; interquartile range [IQR], 1.58–5.38 years) who received isoniazid daily (median dosage, 5.01 mg/kg/day; range, 2.94–15.58 mg/kg/day) as part of antituberculosis treatment. At 1 and 4 months after initiation of treatment, isoniazid concentrations were measured in plasma samples at 0.75, 1.5, 3, 4, and 6 h after a treatment dose, to describe pharmacokinetic measures by using noncompartmental analysis. The effects of dose in milogram per kilogram, acetylator genotype, age, sex, and clinical diagnosis of kwashiorkor and human immunodeficiency virus (HIV) infection on isoniazid concentrations were evaluated.

Results. Median peak concentrations of isoniazid in children prescribed a dose of 4–6 mg/kg were 58% lower than those in children prescribed a dose of 8–10 mg/kg (2.39 mg/L [IQR, 1.59–3.40] vs. 5.71 mg/L [IQR, 4.74–7.62]). Peak concentrations were <3 mg/L in 70% of children prescribed a dose of 4–6 mg/kg. In contrast, children prescribed a dose of 8–12 mg/kg achieved peak concentrations approximating those in adults treated with 300 mg of isoniazid daily. Intermediate or fast acetylator genotype independently predicted a 38% (95% confidence interval [CI], 21%–51%) reduction in peak concentrations, compared with the slow-acetylator genotype. Each 1-mg/kg increase in the dose and each year increase in age were associated with increases in peak concentrations of 21% (95% CI, 16%–25%) and 6% (95% CI, 3%–10%), respectively.

Conclusions. Younger children require higher doses of isoniazid per kilogram of body weight to achieve isoniazid concentrations similar to those in adults. A daily isoniazid dose of 8–12 mg/kg should be recommended.

Isoniazid plays a crucial role in the treatment and prevention of tuberculosis. It has potent bactericidal activity against metabolically active *Mycobacterium tuberculosis* and prevents the development of resistance to companion antituberculosis drugs [1].

Oral doses of isoniazid are rapidly absorbed and readily distributed throughout the body. The drug undergoes intestinal and hepatic first-pass metabolism and is eliminated primarily by acetylation and dehydrazination. The rate of elimination has a trimodal distribution determined by genetic polymorphisms of the arylamine *N*-acetyltransferase 2 gene *(NAT2):* individuals may be homozygous for the allele for a fast acetylator (hereafter, "fast genotype"), heterozygous for the allele for a fast acetylator (hereafter, "intermediate genotype"), or homozygous for the allele for a slow acetylator (hereafter, "slow genotype") [2]. *NAT2* genotype is a prominent determinant of isoniazid concentrations in adults and children [2, 3]. The frequencies of *NAT2*

Received 10 November 2008; accepted 6 January 2009; electronically published 24 April 2009.

Reprints or correspondence: Dr. Helen McIlleron, Div. of Clinical Pharmacology, K45 Old Main Bldg., Groote Schuur Hospital, 7925 Observatory, South Africa (helen.mcilleron@uct.ac.za).

Clinical Infectious Diseases 2009; 48:000–000

1058-4838/2009/4811-00XX$15.00
DOI: 10.1086/598192

polymorphisms vary widely between different populations, and there are corresponding differences in isoniazid pharmacokinetics reported in different regions [3, 4]. Younger children tend to have more-rapid elimination of isoniazid and relatively larger volumes of distribution, in comparison with older children and adults [3, 5–7].

Studies involving adults with pulmonary tuberculosis have demonstrated dose-related efficacy, with regard to both early bactericidal activity [8, 9] and longer-term outcomes [10]. Numerous clinical trials leading to implementation of the 6-month multidrug regimen (rifampin and isoniazid during all 6 months, with pyrazinamide or pyrazinamide and ethambutol during the first 2 months) support a daily isoniazid dose of 300 mg (i.e., 6 mg/kg for a patient weighing 50 kg) for adults [11]. Extensive experience with pharmacokinetic studies involving adults suggests a minimum target of 3 mg/L for the peak concentration (C_{max}) following receipt of a 300-mg dose when isoniazid is administered daily [12]. Reduced isoniazid concentrations may predispose to the development of rifamycin-resistant organisms in patients with tuberculosis who have advanced immunosuppression due to increased periods of exposure to effective concentrations of rifamycin alone [13, 14].

There are multiple obstacles to evaluating antituberculosis treatment regimens and their individual drug components for children [15, 16]. However, it is likely that good responses to treatment will be achieved in children, provided that they are given drug formulations and doses that achieve pharmacokinetics comparable to those that have demonstrated safety and efficacy among adults. Although some national programs, such as those in Japan, Mexico, the Phillipines, and the United States, recommend isoniazid dosages of 10–15 mg/kg/day for children, international treatment guidelines recommend 4–6 mg/kg/day [17, 18].

The South African Tuberculosis Control Programme recommends isoniazid doses of 4–6 mg/kg for children receiving daily doses [19]. However, many pediatricians elect to use higher doses, particularly for children with severe disease. We studied the pharmacokinetics of isoniazid in children prescribed a range of isoniazid doses at the Brooklyn Hospital for Chest Diseases in Cape Town, South Africa.

METHODS

Patients and treatments. Children aged 3 months to 13 years who were referred to the Brooklyn Hospital for Chest Diseases for management of severe forms of tuberculosis were eligible to participate in the study. Parents or legal guardians gave written informed consent before enrollment of their children. Treatment was initiated ~1 month before admission to the Brooklyn Hospital for Chest Diseases by the referring hospital, at which the diagnosis of tuberculosis or probable tuberculosis was made on the basis of ⩾2 of the following criteria: gastric aspirate, sputum, or CSF culture result positive for *M. tuberculosis;* induration >5 mm in children with human immunodeficiency virus (HIV) infection or >10 mm in HIV-uninfected children at 48–72 h after a Mantoux test with 2 U of tuberculin RT23; household source case with sputum-microscopy smear test result positive for acid-fast bacilli in the preceding year; chest radiograph findings indicative of pulmonary tuberculosis; and findings on cranial computed tomography compatible with tuberculous meningitis and appropriate microscopic changes in the cerebrospinal fluid. HIV infection status was determined by enzyme-linked immunosorbent assay, with positive results confirmed by a second assay, or for children aged <18 months, HIV infection status was confirmed by polymerase chain reaction. The parents or legal guardians of the children were counseled and written informed consent was obtained before HIV testing. A clinical diagnosis of kwashiorkor was made in the presence of pitting edema. Children were classified as having fast, intermediate, or slow genotypes for *NAT2* on the basis of previously described methods [3]; the fast and slow alleles are codominant [2]. The study was approved by the Institutional Review Board of the Faculty of Health Sciences of Stellenbosch University (2003/054/N) and was conducted in accordance with national research standards [20] and the Helsinki Declaration of 2000, revised in 2004.

Antituberculosis treatment regimens included daily doses of rifampin and isoniazid for 6 months with pyrazinamide for the first 2 months. Higher doses of isoniazid tended to be prescribed for children with disseminated disease, such as miliary tuberculosis or tuberculous meningitis. Ethionamide was added to the treatment regimen for children with tuberculous meningitis. Ethambutol was added during the intensive phase for other forms of tuberculosis when the use of 4 drugs was considered advisable. Isoniazid was administered as part of dispersible, pediatric, fixed-dose combinations (Rimcure Paed 3-FDC: each tablet contains 60 mg of rifampin, 30 mg of isoniazid, and 150 mg of pyrazinamide; Rimactazid Paed 60/30: each tablet contains 60 mg of rifampin and 30 mg of isoniazid; or Rimactazid Paed 60/60: each tablet contains 60 mg of rifampin and 60 mg of isoniazid; all manufactured by Sandoz). All the antituberculosis medicines were approved by the South African Medicines Control Council and were dispensed by the Brooklyn Hospital for Chest Diseases pharmacy.

Pharmacokinetic assessment. Sample collection for pharmacokinetic analysis was performed at ~1 month after the initial admission to the referral hospital and again at 4 months after initiation of treatment. Children fasted overnight before blood sample collection at 0.75, 1.5, 3.0, 4.0 and 6.0 h after an observed treatment dose. Blood specimens were immediately placed on ice. Plasma (1 mL) was separated by centrifugation within 30 min and was stored at −80°C until analysis. Alanine

aminotransferase (ALT), body mass, and height were measured on the days of pharmacokinetic sample collection.

Plasma concentrations of isoniazid were quantified by tandem high-performance liquid chromatography mass spectrometry (API 2000; Applied Biosystems) with use of a 20 × 2.1–mm Betasil silica column (Thermo). An isocratic elution of 80% acetonitrile in 0.1% formic acid was used as the mobile phase, with a flow rate of 0.3 mL/min and an injection volume of 0.005 mL. The internal standard was sulfamoxole. Selected reaction-monitoring transitions of $[M\text{-}H]^+$ precursor ions to product ions were isoniazid (mass-to-charge ratio [m/z], 138.0–121.2) and sulfamoxole (m/z, 254.0–92.2). Plasma protein was precipitated with 3 volumes of acetonitrile containing the internal standard. Samples were vortexed and centrifuged for 5 min at 750 *g*. Supernatant (0.005 mL) was injected into the column. Standard curves were linear in the range 0.1–15 mg/L. Quality-control samples covering the ranges were included with each run. Interday and intraday coefficients of variation were <10%.

Plasma concentrations at the time of dosing and those >20% below the limit of quantification (0.1 mg/L) were given a value of 0.05 mg/L. WinNonlin, version 4.1 professional (Pharsight), was used to determine C_{max} and time to C_{max} (T_{max}) directly from the concentration-time data; the apparent elimination half-life ($t_{1/2} = 0.693/k_e$, where k_e is the slope of the log-linear regression of ⩾3 final data points); and the area under the curve until the 6-h time point ($AUC_{0–6}$) by the linear trapezoidal rule. $AUC_{0–6}$ and $t_{1/2}$ were not evaluated in patients with missing concentration data at the 4- or 6-h time points.

Statistical analysis. The data were summarized as median and interquartile range (IQR). The Wilcoxon rank-sum test, the Kruskal-Wallis test, and the Wilcoxon signed-rank test were used to determine whether 2 independent groups, 3 independent groups, or paired data, respectively, were statistically significantly different. For binominal data, differences between groups were determined using Fisher's exact test. The Spearman rank correlation coefficient described associations between continuous variables. The natural logarithm transformation was applied to C_{max} values for examination of the covariate effects of age, sex, HIV infection status, a diagnosis of kwashiorkor, isoniazid dose per kilogram of body weight, and *NAT2* genotype (as a binary variable: slow genotype in one category; intermediate and fast genotypes in the other category) by use of linear regression. Effects of covariates that were found to be possibly associated ($P < .2$) with the log-transformed C_{max} in the univariate analyses were further examined by multivariate regression analysis. The regression coefficients were back-transformed to express the percentage change in the dependent variable conferred by a 1-unit change in the relevant covariate. The model assumptions of constant variance, linearity, and the appropriate form of the covariates in the model were checked using methods based on the distribution of the residuals. The odds of developing transaminitis was evaluated for children treated with an isoniazid dose of 8–12 mg/kg, compared with children treated with a dose of 4–6 mg/kg, in a univariate logistic regression model. Stata, version 8.2 (StataCorp), was used to compute summary statistics and for statistical tests and regression analysis.

RESULTS

Patient characteristics and treatment doses. A total of 60 children were enrolled in the study. Four HIV-infected children were withdrawn before pharmacokinetic assessment because they required transfer to other hospitals for management of complications. Of the remaining 56 children, 29 (52%) were male, and 22 (39%) were infected with HIV-1. Of the 56 children, 20 (36%) had slow genotypes, 24 (43%) had intermediate genotypes, and 8 (14%) had fast genotypes. (Genotyping results were not available for 4 participants with a mean age of 2.76 years, mean weight of 11.3 kg, and mean height of 82 cm). Their median age was 3.22 years (IQR, 1.58–5.38 years), median weight was 12.43 kg (IQR, 8.88–17.26 kg), and median height was 84.5 cm (IQR, 76.2–104.2 cm) at 1 month after initiation of antituberculosis treatment. The dose of isoniazid at the first pharmacokinetic assessment ranged from 2.94 to 15.58 mg/kg (median, 5.01 mg/kg; IQR, 4.35–9.24 mg/kg) or 52 to 317 mg/m^2 of body surface area (median, 119 mg/m^2; IQR, 100–198 mg/m^2). Infants (age, <1 year) were more likely to receive isoniazid doses of <4 or >12 mg/kg (3 of 6 infants vs. 6 of 50 children aged >1 year; $P = .046$, by Fisher's exact test), and children who received a diagnosis of kwashiorkor tended to be prescribed higher doses than those without a diagnosis of kwashiorkor (median, 6.27 mg/kg [IQR, 4.90–9.72 mg/kg] vs. 4.82 mg/kg [IQR, 4.25–8.77 mg/kg]; $P = .074$, by Wilcoxon rank-sum test). Doses were not statistically significantly different between children with and children without HIV infection (median, 4.61 mg/kg [IQR, 4.23–7.33 mg/kg] vs. 5.14 mg/kg [IQR, 4.72–9.27 mg/kg]; $P = .191$, by Wilcoxon rank-sum test), between female and male patients (median, 5.23 mg/kg [IQR, 4.50–9.20 mg/kg] vs. 4.85 mg/kg [IQR, 4.25–9.27 mg/kg]; $P = .372$, by Wilcoxon rank-sum test), or among children with slow, intermediate, and fast genotypes (table 1).

Pharmacokinetics. Isoniazid concentrations at ~1 month after initiation of antituberculosis treatment were available for 56 participants. Isoniazid concentrations were not available for 4 and 13 samples obtained at the 4- and 6-h time points, respectively, because of our inability to maintain the intravenous line for the duration of sample collection or because of an insufficient sample for analysis of isoniazid concentrations. Two samples obtained at the 4-h time point and 3 samples obtained at the 6-h time point were below the limit of quantification. Plasma concentrations at each sample collection time

varied widely among individuals (figure 1). Absorption was rapid, with a median T_{max} of 0.75 h. The dose, C_{max}, $t_{1/2}$, AUC_{0-6}, and concentration at the 6-h time point (C_6) for children with slow, intermediate, and fast *NAT2* genotypes are reported in table 1, and C_{max} and AUC_{0-6} for children receiving doses <4 mg/kg, 4–6 mg/kg, >6 and <8 mg/kg, 8–12 mg/kg, and >12 mg/kg are described in table 2. C_{max} and AUC_{0-6} were highly correlated (Spearman rank correlation coefficient, 0.938; $P < .001$).

None of the 7 children (3 with slow genotypes and 4 with intermediate genotypes) who received isoniazid doses <4 mg/kg had a C_{max} above the recommended lower limit of 3 mg/L. Of 30 children prescribed doses of 4–6 mg/kg, 21 (70%) had a C_{max} <3 mg/L (4 [36%] of 11 with slow genotypes, 9 [90%] of 10 with intermediate genotypes, 6 [100%] of 6 with fast genotypes, and 2 [67%] of 3 with unknown genotypes). All children prescribed doses of 8–12 mg/kg had peak concentrations >3 mg/L (5 children with slow genotypes and median C_{max} = 7.10 mg/L, 8 children with intermediate genotypes and median C_{max} = 5.35 mg/L, 1 child with fast genotype and C_{max} = 5.58 mg/L, and 1 child with unknown genotype and C_{max} = 5.70 mg/L).

Univariate regression analyses found no association ($P > .2$) between isoniazid C_{max} at 1 month after initiation of antituberculosis treatment and HIV infection status or between C_{max} at 1 month and the presence of kwashiorkor. Sex was weakly associated ($P = .104$) in univariate analysis but did not contribute to the multivariate model. A multivariate model describing 71% of the variability in C_{max} found that each 1-mg/kg increase in dose was associated with a 21% increase (95% CI, 16%–25% increase) in C_{max} and that each 1-year increase in age was associated with a 6% increase (95% CI, 3%–10% increase) in C_{max}. Participants with intermediate or fast *NAT2* genotypes had 38% decreases (95% CI, 21%–51% decrease) in C_{max}, compared with those with slow genotypes. The model included 51 observations, because *NAT2* genotype was unknown for 4 participants and because there was 1 outlying observation (C_{max} = 20.08 mg/L; dose, 5.06 mg/kg; slow genotype; age, 2.41 years) that was excluded to satisfy the math-

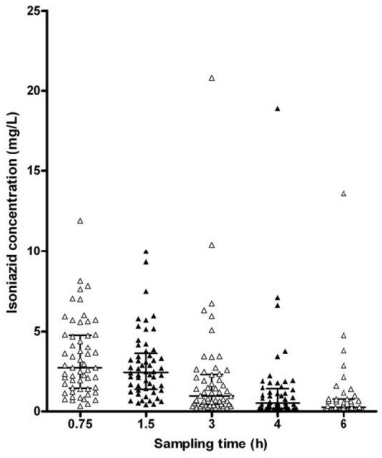


Figure 1. Isoniazid concentrations at each sample collection time in 56 children with tuberculosis at 1 month after initiation of antituberculosis treatment. *Middle bars,* medians; *bars connected by horizonal line,* interquartile ranges.

Table 1. Pharmacokinetic assessment at 1 month after initiation of antituberculosis treatment, according to arylamine *N*-acetyltransferase 2 (*NAT2*) genotype.

Isoniazid measure	All patients[a]	Children with *NAT2* genotype			P[b]
		Slow[c]	Intermediate[d]	Fast[e]	
Dose, mg/kg	5.01 (4.35–9.24)	4.98 (4.12–7.99)	5.26 (4.43–9.37)	4.90 (4.68–7.38)	.748
C_{max}, mg/L	3.07 (1.82–5.66)	4.05 (2.72–5.74)	2.63 (1.61–5.25)	1.54 (1.22–4.14)	.067
$t_{1/2}$, h	1.59 (1.21–2.17)	2.23 (1.80–3.06)	1.36 (1.14–1.75)	1.12 (0.96–1.26)	<.001
AUC_{0-6}, mg·h/L	8.73 (4.55–14.13)	10.60 (8.66–16.14)	6.45 (4.18–13.27)	2.31 (1.77–6.12)	.014
C_6, mg/L	0.26 (0.12–0.76)	0.76 (0.52–1.86)	0.14 (0.11–0.26)	0.08 (0.05–0.12)	<.001

NOTE. Data are median (interquartile range), unless otherwise indicated. *NAT2* genotype was available for 52 of the 56 children in the study. AUC_{0-6}, area under the curve until the 6-h time point; C_{max}, peak concentration; C_6, concentration at the 6-h time point; $t_{1/2}$, half life.

[a] For dose and C_{max}, $n = 56$; for $t_{1/2}$, $n = 40$; for AUC_{0-6}, $n = 41$; and for C_6, $n = 43$. AUC_{0-6} and $t_{1/2}$ were not determined for patients with missing samples at the 4- or 6-h time point. For 1 patient with a late C_{max} measured at 4 h after the treatment dose, $t_{1/2}$ was not estimated.

[b] By Kruskal-Wallis test to determine the probability of equality between the genotype groups.

[c] For dose and C_{max}, $n = 20$; for $t_{1/2}$, $n = 14$; for AUC_{0-6}, $n = 15$; and for C_6, $n = 16$.

[d] For dose and C_{max}, $n = 24$; for $t_{1/2}$ and AUC_{0-6}, $n = 20$; and for C_6, $n = 21$.

[e] For dose and C_{max}, $n = 8$; and for $t_{1/2}$, AUC_{0-6}, and C_6, $n = 4$.

Table 2. Pharmacokinetics at 1 month after initiation of antituberculosis treatment and patient characteristics, according to strata of isoniazid dose.

Isoniazid dose stratum, mg/kg	No. of children	Isoniazid dose, mg/kg	Age at enrollment, years	C_{max}, mg/L	AUC_{0-6}, mg·h/L[a]
<4	7	3.55 (3.25–3.82)	1.27 (0.41–5.10)	0.76 (0.69–2.24)	2.93 (1.56–7.20)
4–6	30	4.77 (4.37–5.08)	3.58 (1.97–5.51)	2.39 (1.59–3.40)	5.97 (4.00–9.39)
>6 to <8	2	7.00 (6.66–7.33)	7.01 (2.04–11.97)	5.85 (5.70–6.00)	11.72 (11.09–12.35)
8–12	15	9.72 (9.27–10.81)	4.07 (1.91–6.43)	5.71 (4.74–7.62)	14.13 (9.40–28.42)
>12	2	14.39 (13.19–15.58)	0.81 (0.61–1.01)	6.46 (5.92–6.99)	19.74

NOTE. Data are median (interquartile range), unless otherwise indicated. AUC_{0-6}, area under the curve until the 6-h time point; C_{max}, peak concentration.

[a] AUC_{0-6} was determined for 41 children: for dose stratum <4 mg/kg, $n = 3$; for dose stratum 4–6 mg/kg, $n = 24$; for dose stratum >6 to <8 mg/kg, $n = 2$; for dose stratum 8–12 mg/kg, $n = 11$; and for dose stratum >12 mg/kg, $n = 1$.

ematical assumptions of the model. Substitution of dose in mg/kg and age with dose in mg/m^2 body surface area resulted in a simplified multivariate model that described the data equally well ($r^2 = 71\%$; intermediate and fast genotypes were associated with a 36% lower [95% CI, 19%–49% lower] C_{max}, compared with slow genotypes; and each 25-mg/m^2 increase in the isoniazid dose resulted in a 24% increase [95% CI, 20%–31% increase] in C_{max}).

One child with HIV infection and 1 child without HIV infection were discharged from the hospital after completion of the first pharmacokinetic evaluation. The remaining 54 children underwent additional pharmacokinetic evaluation at ~4 months after initiation of antituberculosis treatment. Children had statistically significant growth during the intervening 3 months; the median weight gain was 0.48 kg (IQR, −1.00 to 1.35 kg; $P < .001$, by Wilcoxon signed-rank test), and the median increase in height was 1.75 cm (IQR, 0.55–3.00 cm; $P < .001$, by Wilcoxon signed-rank test). Isoniazid doses were adjusted to keep pace with weight gain; therefore, doses in mg/kg were not significantly different at 1 and 4 months (median dose change, −0.05 mg/kg; IQR, −0.46 to 0.18 mg/kg; $P = .384$, by Wilcoxon signed-rank test). The pharmacokinetics of isoniazid at 1 and 4 months after initiation of antituberculosis treatment were similar. At 4 months, isoniazid concentrations were not available for 3 and 10 samples obtained at the 4- and 6-h time points, respectively. One and 3 samples obtained at the 4- and 6-h time points, respectively, were below the limit of quantification. Median changes in pharmacokinetic variables were as follows: C_{max}, 0.10 mg/L (IQR, −0.79 to 1.18 mg/L; $P = .446$, by Wilcoxon signed-rank test); C_6, 0.02 mg/L (IQR, −0.05 to 0.09 mg/L; $P = .479$, by Wilcoxon signed-rank test); AUC_{0-6}, 1.39 mg·h/L (IQR, −1.40 to 3.48 mg·h/L; $P = .343$, by Wilcoxon signed-rank test), and $t_{1/2}$, −0.02 h (IQR, −0.39 to 0.24 h; $P = .550$, by Wilcoxon signed-rank test). Multivariate regression confirmed the important effects of dose (mg/kg) and *NAT2* genotype at 4 months on the systemic concentrations of isoniazid. Age, sex, HIV infection, and kwashiorkor were not associated with C_{max} after adjustment for the 2 dominant covariates.

Serum ALT level measurements were unavailable for 2 children at 1 month and for 3 children at 4 months after initiation of antituberculosis treatment. Two children (both female with intermediate *NAT2* genotype; one was treated with isoniazid at a dosage of 2.94 mg/kg/day, and the other, who had kwashiorkor and HIV infection, was treated with isoniazid at a dosage of 13.19 mg/kg/day) had ALT levels >5 times the upper limit of normal at 1 month after initiation of antituberculosis treatment. In both cases, ALT was within the normal range at 4 months after intiation of antituberculosis treatment. One child (a female with intermediate genotype and HIV infection who was treated with isoniazid at a dosage of 4.23 mg/kg/day) had an ALT level >5 times the upper limit of normal at 4 months. The risk of developing transaminitis (ALT level, >1.25 times the upper limit of normal) at 1 month (6 children) was not significantly greater for children treated with isoniazid doses of 8–12 mg/kg, compared with those treated with doses of 4–6 mg/kg (OR, 2.00; 95% CI, 0.25–15.85).

DISCUSSION

We describe the pharmacokinetics of isoniazid and *NAT2* genotypes in a cohort of 56 children with severe forms of tuberculosis. Although previous studies have characterized the concentrations of isoniazid in South African children treated with higher doses [3], the pharmacokinetics of the 4–6-mg/kg dose that is recommended by international authorities, including the World Health Organization, has not been described in South African children. Peak concentrations were below the recommended reference range in 70% of children (21 of 30, including 4 [36%] of 11 children with slow genotypes) who were prescribed isoniazid doses of 4–6 mg/kg at 1 month after initiation of antituberculosis treatment. Thus, even among populations with a high prevalence of slow genotypes, a substantial proportion of children may achieve relatively low isoniazid con-

centrations. In contrast, all children prescribed isoniazid doses of 8–12 mg/kg, irrespective of *NAT2* genotype, achieved a $C_{max} > 3$ mg/L, with a median C_{max} comparable to that reported in South African adults (6.5 mg/L; IQR, 4.9–8.7 mg/L) who were treated with 300 mg of isoniazid daily [21].

Children prescribed doses <4 mg/kg had low isoniazid concentrations, and infants were at greater risk than were older children of receiving low or high doses. Thus, the importance of accurate weighing of young children should be emphasized, as should the need to increase doses in line with growth and weight gain during treatment. It is relevant to note that the South African National Tuberculosis Control Programme and the package inserts for the fixed-dose combination formulations used in this study recommend that children who weigh 3–4.9 kg should be prescribed one-half of a fixed-dose combination tablet containing 30 mg of isoniazid [19, 22–24]. Thus, a child with a weight of 4.9 kg would receive only 3.06 mg/kg of isoniazid if the tablet was accurately halved. Likewise, 1 tablet is recommended for children who weigh 5–7 kg, and a dose of 3.80 mg/kg is recommended for a child who weighs 7.9 kg.

Dose in mg/kg of body weight and *NAT2* genotype were the dominant determinants of isoniazid concentrations. After adjustment for the effects of dose and *NAT2* genotype, age was associated with isoniazid concentrations at ~1 month after initation of treatment. Age was noncontributory to the model describing the effects of genotype and dose in mg/m^2 body surface area on C_{max}, which suggests that dosing based on body surface area may result in less variability in isoniazid concentrations among young children than might dosing based on weight. Age did not improve the model describing the effects of genotype and dose in mg/kg on C_{max} after 4 months of treatment. The fact that the children were older might explain in part the lack of association; however, the multivariate analysis at 4 months described only 48% of the variability in C_{max}, compared with 71% in the earlier analysis indicating greater residual variability unexplained by the covariates investigated. Nonetheless, this finding highlights the need for further evaluation of isoniazid concentrations in infants (only 6 children included in this study were aged <1 year). The similarity in isoniazid pharmacokinetics at 1 and 4 months after initation of antituberculosis treatment suggests that any physiological changes affecting isoniazid concentrations related to disease severity in this group of children had resolved within 1 month of treatment. Thus, the pharmacokinetic results are likely to reflect results for children with milder disease.

Data relating to treatment outcomes were not available. Although this is an important limitation, the assessment of response to antituberculosis treatment in a standardized manner is difficult in children because reliable markers of treatment response have not been developed [15, 16].

Severe isoniazid-related hepatitis is rare in children [25]. Although the risk of transaminitis tended to be increased in those receiving higher doses of isoniazid, this finding was not statistically significant and may have been confounded by disease-related factors, because children with disseminated tuberculosis tended to be prescribed higher doses. All children received a multivitamin supplement that supplied (at least) the recommended daily allowance of pyridoxine. None of the cohort developed peripheral neuropathy. It is our clinical experience (and that of others [26, 27]) that peripheral neuropathy is very rare in HIV-uninfected children treated with isoniazid, even at doses of up to 20 mg/kg. However, pyridoxine supplementation is prudent for children with HIV infection or nutritional deficiency.

In conclusion, in this group of young children, a dose of 4–6 mg/kg was insufficient to achieve isoniazid concentrations comparable to those deemed necessary for optimal response in adults. Conversely, a dose of 8–12 mg/kg achieved comparable peak concentrations to a 300-mg dose in adults.

Acknowledgments

We thank the staff at the Brooklyn Hospital for Chest Diseases and the Clinical Pharmacology Laboratory, University of Cape Town, for their contributions, and Karien Cilliers, for her assistance with anthropometric data.

Financial support. Bristol-Myers Squibb "Secure the Future" Foundation.

Potential conflicts of interest. H.M. has received an honorarium from Abbott Laboratories. P.R.D. has been employed as a consultant by Tibotec Pharmaceuticals. All other authors: no conflicts.

References

1. Mitchison DA. Role of individual drugs in the chemotherapy of tuberculosis. Int J Tuberc Lung Dis **2000**; 4:796–806.
2. Parkin DP, Vandenplas S, Botha FJH, et al. Trimodality of isoniazid elimination. Am J Respir Crit Care Med **1997**; 155:1717–22.
3. Schaaf HS, Parkin DP, Seifart HI, et al. Isoniazid pharmacokinetics in children treated for respiratory tuberculosis. Arch Dis Child **2005**; 90: 614–8.
4. Roy V, Tekur U, Chopra K. Pharmacokinetics of isoniazid in pulmonary tuberculosis—a comparative study at two dose levels. Indian Pediatr **1996**; 33:287–91.
5. Lupasco I, Algeorge G, Ghiu-Cimpeanu V. Taux d'isoniazide actif dans le sérum des enfants tuberculeux. Poumon Coeur **1965**; 21:1105–14.
6. Advenier C, Saint-Aubin A, Scheinmann P, Paupe J. Pharmacocinétique de l'isoniazide chez l'enfant. Rev Fr Mal Respir **1981**; 9:365–74.
7. Bouveret JP, Hanoteau J, Gerbeaux J, Houin G, Tillement JP. Variations de l'indice d'inactivation de l'isoniazide au cours des traitements antituberculeaux chez l'enfant. Arch Fr Pediatr **1983**; 40:615–9.
8. Donald PR, Sirgel FA, Botha FJH, et al. The early bactericidal activity of isoniazid related to its dose size in pulmonary tuberculosis. Am J Respir Crit Care Med **1997**; 156:895–900.
9. Donald PR, Sirgel FA, Venter A, et al. The influence of human N-acetyltransferase genotype on the early bactericidal activity of isoniazid. Clin Infect Dis **2004**; 39:1425–30.
10. Tuberculosis Chemotherapy Centre, Madras. A concurrent comparison of isoniazid plus PAS with three regimens of isoniazid alone in the domiciliary treatment of pulmonary tuberculosis in South India. Bull World Health Organ **1960**; 23:535–85.
11. Fox W, Ellard GA, Mitchison DA. Studies on the treatment of tuberculosis undertaken by the British Medical Research Council tubercu-

losis units, 1946–1986, with relevant subsequent publications. Int J Tuberc Lung Dis **1999**; 3:S231–79.
12. Peloquin CA. Therapeutic drug monitoring in the treatment of tuberculosis. Drugs **2002**; 62:2169–83.
13. Weiner M, Burman W, Vernon A, et al. Low isoniazid concentrations and outcome of tuberculosis treatment with once-weekly isoniazid and rifapentine. Am J Respir Crit Care Med **2003**; 167:1341–7.
14. Wallis RS, Weyer K, Fourie PB. Acquired rifamycin resistance: pharmacology and biology. Expert Rev Anti Infect Ther **2008**; 6:223–30.
15. Donald PR. The assessment of new anti-tuberculosis drugs for a paediatric indication. Int J Tuberc Lung Dis **2007**; 11:1162–5.
16. Burman WJ, Cotton MF, Gibb DM, Walker AS, Vernon AA, Donald PR. Ensuring the involvement of children in the evaluation of new tuberculosis treatment regimens. PloS Med **2008**; 5:e176.
17. World Health Organization. Guidance for national tuberculosis programmes on the management of tuberculosis in children. Document WHO/HTM/TB/2006.371. Geneva: World Health Organization, **2006**. Available at: http://www.who.int/tb/publications/2006/en/index.html. Accessed 24 October 2008.
18. Tuberculosis Coalition for Technical Assistance. International standards for tuberculosis care. The Hague: Tuberculosis Coalition for Technical Assistance, **2006**. Available at: http://www.tbcta.org/TB_CAP_Toolbox/#15. Accessed 24 October 2008.
19. National Tuberculosis Control Programme practical guidelines 2000. Pretoria, South Africa: Department of Health, Republic of South Africa, **2000**. Available at: http://www.doh.gov.za/tb/index.html (under Resources: Basic documents and reports). Accessed 24 October 2008.
20. Guidelines for good practice in the conduct of clinical trials in human participants in South Africa. Pretoria, South Africa: Department of Health, Republic of South Africa, **2000**. Available at: http://www.doh.gov.za/docs/policy/trials/trials-full.html. Accessed 24 October 2008.
21. McIlleron H, Wash P, Burger A, Norman J, Folb PI, Smith P. Determinants of rifampin, isoniazid, pyrazinamide, and ethambutol pharmacokinetics in a cohort of tuberculosis patients. Antimicrob Agents Chemother **2006**; 50:1170–7.
22. Rimactazid Paed 60/60 tablets (rifampicin and isoniazid) [package insert]. Kempton Park, South Africa: Rolab, **1998**.
23. Rimactazid Paed 60/30 tablets (rifampicin and isoniazid) [package insert]. Kempton Park, South Africa: Rolab, **1998**.
24. Rimcure Paed 3-FDC tablets (rifampicin, isoniazid, and pyrazinamide) [package insert]. Kempton Park, South Africa: Rolab, **1998**.
25. Wu SS, Chao CS, Vargas JH, et al. Isoniazid-related hepatic failure in children: a survey of liver transplantation centers. Transplantation **2007**; 84:173–9.
26. Morales SM, Lincoln EM. The effect of isoniazid therapy on pyridoxine metabolism in children. Am Rev Tuberc **1957**; 75:594–600.
27. Miller FJW, Seal RME, Taylor MD. Tuberculosis in children. London: J & A Churchill, **1963**:163–87.

CHAPTER 11

The Influence of Dose and *N*-acetyltransferase-2 (*NAT2)* Genotype and Phenotype on the Pharmacokinetics and Pharmacodynamics of Isoniazid

The influence of dose and *N*-acetyltransferase-2 (NAT2) genotype and phenotype on the pharmacokinetics and pharmacodynamics of isoniazid

P.R. Donald[1§], D.P. Parkin[2], H.I. Seifart[2], H.S. Schaaf[1], P.D. van Helden[3], C.J. Werely[3], F.A. Sirgel[3], A. Venter[3] and J.S. Maritz[4]

[1]Department of Paediatrics and Child Health, Faculty of Health Sciences, Stellenbosch University, P.O. Box 19063, Tygerberg 7505, South Africa

[2]Department of Pharmacology, Faculty of Health Sciences, Stellenbosch University, P.O. Box 19063, Tygerberg 7505, South Africa.
[3]Department of Biology and Human Genetics, and the MRC Centre for Molecular and Cellular Biology, DST/NRF Centre of Excellence for Biomedical TB Research, Faculty of Health Sciences, Stellenbosch University, P.O. Box 19063, Tygerberg 7505, South Africa.
[4]Biostatistical Unit of the South African, P.O. Box 19063, Tygerberg 7505, South Africa.

[§]Corresponding Author (email: prd@sun.ac.za)

European Journal of Clinical Pharmacology. (2007) **63:**633-639.

Précis

INH is still one of the most effective agents used in anti-tuberculosis treatment despite being one of the oldest. It targets actively growing mycobacterial cells, reducing their numbers within a very short time period, referred to as its "early bactericidal activity" (EBA). It has previously been shown that the *NAT2* genotype has a significant effect on the EBA of INH, and therefore we assayed INH dosage to investigate an optimal EBA, in which a "killing power" of 90% or greater (EBA_{90}) could be obtained.

A cohort of TB patients, the majority of whom were from the South African Coloured (SAC) group, were given a range of INH doses, from 0.2 mg/kg to 12 mg/kg body weight. The acetylation status of the patients was assessed via *NAT2* genotype analysis, as well as the optimum EBA at each INH dose. We found that the optimum EBA (EBA_{90}) was attained at different doses of INH in the fast (FF), intermediate (FS) and slow (SS) acetylation groups. These results indicated that, whereas a low dose of INH in slow acetylators would be sufficient to attain optimum treatment efficacy, this level of INH would be sub-optimal in intermediate and fast acetylators. Furthermore, the percentage of patients that attained the EBA_{90} increased proportionately, in all three acetylation groups, with increasing INH dosage, such that at a dose of 10-12 mg/kg INH, all of the patients attained the EBA_{90}.

The results of this study, whilst promising, only relate to the early effects of INH in the different acetylator groups, and it would be interesting to determine if this early effect of INH also applies to the end of therapy (treatment outcome). Clearly these results for INH have to be interpreted in conjunction with co-administered drugs under standardised TB therapy. Since the 10-12 mg/kg dose is a higher dose of INH than generally used, it is likely that more cases of hepatotoxicity (especially among SS acetylators) may occur, and this will need to be investigated in future.

Sample Source:- Adult patient samples were obtained from Tygerberg Hospital and the Brooklyn Hospital for Chest Diseases. Sample cohort used in this study only.

All NAT2 genotyping was done by CJW, who also contributed the interpretation and writing of the manuscript.

Eur J Clin Pharmacol
DOI 10.1007/s00228-007-0305-5

PHARMACOKINETICS AND DISPOSITION

The influence of dose and *N*-acetyltransferase-2 (NAT2) genotype and phenotype on the pharmacokinetics and pharmacodynamics of isoniazid

P. R. Donald · D. P. Parkin · H. I. Seifart · H. S. Schaaf · P. D. van Helden · C. J. Werely · F. A. Sirgel · A. Venter · J. S. Maritz

Received: 10 January 2007 / Accepted: 22 March 2007

Abstract
Objective This study evaluated the pharmacokinetics of isoniazid (INH) associated with optimal early bactericidal activity (EBA), defined as 90% of the maximum EBA (EBA_{90}) and the influence of *N*-acetyltransferase-2 (NAT2) subtype on the ability of pulmonary tuberculosis (PTB) patients to reach the identified pharmacokinetic values after INH doses ranging from 0.2 to 10–12 mg/kg body weight.
Methods INH serum concentrations and NAT2 subtype were determined during four studies of PTB patients in three of whom the EBA of INH was determined. The relationship of EBA to area under the curve (AUC) ($AUC_{0-\infty}$) and 2-h serum concentrations was examined by exponential regression and fitted curves estimated the $AUC_{0-\infty}$ and 2-h serum concentrations at which EBA_{90} was reached.
Results EBA_{90} was reached at an $AUC_{0-\infty}$ of 10.52 μg/ml per hour and 2-h serum concentrations of 2.19 μg/ml. An $AUC_{0-\infty}$ of 10.52 μg/ml per hour was reached by all 66 patients receiving a 10–12 mg/kg INH dose and all 21 receiving 6 mg/kg, except 1 of 10 (10%) homozygous fast (FF) acetylators; however, at 5 mg/kg, 4 of 12 (33%) FF and 26 of 27 (96%) heterozygous fast (FS), but all 21 homozygous slow (SS) acetylators did so; and 1 of 3 (33%) FF, 2 of 6 (33%) FS, but all 4 SS acetylators at dose 3 mg/kg. An INH 2-h serum concentration of 2.19 μg/ml was reached by all 66 patients receiving 10–12 mg/kg and all 21 receiving 6 mg/kg, except for 2 (20%) FF acetylators at a dose of 5 mg/kg; however, only 3 (25%) of 12 FF acetylators, but 26 (96%) of 27 FS acetylators, and all 21 SS acetylators reached this concentration; and at a dose of 3 mg/kg, 1 (33%) of 3 FF acetylators, 2 (33%) of 6 FF, but all 4 SS acetylators.
Conclusions At a 6 mg/kg dose, all except a minority of FF NAT2 acetylators, achieve an INH $AUC_{0-\infty}$ and 2-h INH serum concentrations associated with EBA_{90}, as did all 4 SS acetylators receiving 3 mg/kg. Any dose reduction below 6 mg/kg body weight will tend to disadvantage a significant proportion of faster acetylators, but, conversely, SS acetylators require only a 3 mg/kg dose to achieve a satisfactory exposure to INH.

Keywords Isoniazid · Early bactericidal activity · *N*-acetyltransferase-2 · Pharmacokinetics · Pharmacodynamics

P. R. Donald (✉) · H. S. Schaaf
Paediatrics and Child Health, Faculty of Health Sciences,
Stellenbosch University,
P.O. Box 19063, Tygerberg 7505, South Africa
e-mail: prd@sun.ac.za

D. P. Parkin · H. I. Seifart
Department of Pharmacology, Faculty of Health Sciences,
Stellenbosch University,
P.O. Box 19063, Tygerberg 7505, South Africa

P. D. van Helden · C. J. Werely · F. A. Sirgel · A. Venter
Department of Biology and Human Genetics,
and the MRC Centre for Molecular and Cellular Biology,
DST/NRF Centre of Excellence for Biomedical TB Research,
Faculty of Health Sciences, Stellenbosch University,
P.O. Box 19063, Tygerberg 7505, South Africa

J. S. Maritz
Biostatistics Unit of the South African Medical Research Council,
Francie van Zyl Drive,
Tygerberg 7505, South Africa

Introduction

Isoniazid (INH) is one of the oldest, but still one of the most effective, antituberculosis agents. Introduced into

clinical practice in 1952 [14, 21, 31], it has an unrivaled early bactericidal activity (EBA), or ability to kill the rapidly metabolizing bacilli found in the sputum of microscopy smear-positive pulmonary tuberculosis patients during the first 2 days of treatment [6, 9, 19]. Perhaps because of this property, it also remains the most effective agent for the prevention of resistance in companion drugs [24].

INH is eliminated by the polymorphic *N*-acetyltransferase-2 (NAT2) enzyme system of the liver and small intestine. Individuals can be partitioned into three genotypes, homozygous slow (SS), heterozygous fast (FS), and homozygous fast (FF), as regards their ability to eliminate INH, and the exposure of the different genotypes to INH differs considerably [28]. None the less, these differences have been considered unimportant amongst compliant patients receiving daily or three or two times weekly intermittent multidrug antituberculosis therapy [13]. Recently, however, the deficiencies to which FF acetylators of INH are exposed were again demonstrated by the development of rifamycin monoresistance in HIV-infected individuals receiving once-weekly intermittent therapy [38]. In another study, acquired rifamycin resistance was associated with low INH area under the curve (AUC) [39]. In an earlier study, we showed that the EBA of INH was significantly influenced by the NAT2 genotype and that a maximum EBA was reached between INH serum concentrations of 2–3 μg/ml [9].

For the purpose of this paper, we reanalyzed INH EBA data obtained during three studies of INH and other antituberculosis drugs [6, 8, 9] together with simultaneously determined INH serum concentrations to determine more precisely what serum concentrations of INH are associated with the optimum killing of metabolically active, fully drug sensitive, *Mycobacterium tuberculosis* bacilli. For this purpose, we defined an optimal EBA as that which is 90% or more of the maximum EBA achieved, or the EBA_{90}. We also reexamined INH pharmacokinetic data from the above studies and, in addition, pharmacokinetic data from a group of adults with pulmonary tuberculosis [28] to evaluate what proportion of SS, FS, and FF NAT2 acetylators of INH receiving a spectrum of INH doses would reach these standards.

Patients and methods

The data regarding the INH serum concentrations analyzed in this paper are drawn from four studies during which a spectrum of INH doses was given to groups of adults and in whom the NAT2 genotype was also ascertained. During three of these studies, the EBA of INH was also determined.

1. Sixty pulmonary tuberculosis patients were recruited and randomly allocated to two treatment sequences [28]. In the first sequence, the patient received INH 5 mg/kg orally on day 1 and 10 mg/kg on day 2; and in the second, sequence the dosing order was reversed. A baseline blood specimen was drawn immediately before INH administration and 0.2, 1, 2, 3, 4.5, and 6 h thereafter. NAT2 genotype was ascertained in 47 patients, and in the remaining 13, the NAT2 phenotype could be assigned with confidence from the pharmacokinetic data [28].
2. Patients were recruited for a study of the EBA of INH after doses ranging from 0.2 mg/kg to 12 mg/kg [6]. Blood specimens were drawn for the determination of INH concentrations at 2, 3, 4, and 5 h after INH administration from nine patients, three receiving a 0.2 mg/kg dose of INH and six receiving 12 mg/kg. The NAT2 genotype was determined in all nine patients.
3. During a study to evaluate the influence of the NAT 2 genotype on the EBA of INH, patients were randomized to receive INH in doses of 1.5 mg/kg, 3 mg/kg, or 6 mg/kg [9]. Blood for the determination of INH serum concentrations was drawn 2, 3, 4, and 5 h after INH administration from 13 patients receiving 6 mg/kg per body weight INH, 18 receiving 3 mg/kg, and 18 receiving 1.5 mg/kg. The NAT2 genotype was determined in 41 of these patients.
4. For a study of the EBA of streptomycin, a control group of 12 patients was randomized to receive 6 mg/kg of INH [8]. In 11 of these patients, INH serum concentrations were determined at 2, 3, 4, and 5 h after drug administration, and the NAT2 genotype was ascertained in nine.

Table 1 summarizes the characteristics of the patients enrolled in these studies. Patients were eligible for enrollment only if they suffered from uncomplicated sputum smear-positive pulmonary tuberculosis. Patients were thus not receiving any other agents at the time of the studies. HIV testing was not carried out on all patients, but in studies in which this was done, only approximately 5% of patients were HIV infected [7]. The great majority of patients were of mixed race or colored racial group.

Early bactericidal activity The EBA of an antituberculosis agent is defined as the fall in $\log_{10}$ colony forming units (CFU) of *M. tuberculosis* in the sputum of patients with microscopy smear-positive pulmonary tuberculosis during the first 2 days of treatment [19]. Sputum specimens for the determination of EBA were collected overnight for 16 h in wide-mouthed sputum jars and delivered as soon as possible to the laboratory for processing. Specimens that could not be delivered immediately to the laboratory were

Table 1 Demographic and clinical features of patients

Patient features									
Study	*n*	Mean age (years)	Males (%)	Mean weight (kg)	Multicavitary disease (%)	NAT2 subtype			
						n	SS	FS	FF
1	60	34.5	72	51.7	–	60	21	27	12
2	10	29.0	77	51.0	68	9	3	2	4
3	49	31.1	69	51.3	98	41	17	15	9
4	11	36.1	73	52.1	100	9	1	3	5

NAT2 N-acetyltransferase-2, *SS* homozygous slow NAT2 genotype, *FS* heterozygous NAT2 genotype, *FF* homozygous fast NAT2 genotype

placed on ice and reached the laboratory within 48 h. The first sputum specimen (S1) was collected before the start of treatment. Following 2 days of treatment, the S3 specimen was collected. If a sputum specimen appeared diluted by saliva or was bloodstained, it was not processed. The specimens were examined by conventional smear, culture, and species identification and for sensitivity to INH. On arrival in the laboratory, the sputum volume was measured and the specimen homogenized by stirring with a Teflon-coated magnetic follower rod for 30 min. Ten milliliters of the homogenate was added to an equal volume of dithiothreitol (Sputolysin, Hoechst) in a 50-ml screw-top tube containing three to six glass beads and was vortex mixed for 20 s. The specimen was then mixed for a further 30 min by mechanical shaking. Two series of tenfold dilutions were prepared in distilled water. From each dilution, aliquots of 100 µl were spread onto half plates of 7H11 oleic acid albumin agar medium made selective by the addition of Selectatabs (Mast, Bootle, UK) to give final concentrations of polymyxin B sulphate 200 units/ml, carbenicillin 100 µg/ml, trimethoprim 10 µg/ml, and amphoteracin B 10 µg/ml. These plates were placed in a polythene bag together with a plate inoculated with *M phlei* to provide CO_2 and incubated for 3 weeks. CFU were counted at that dilution, permitting the counting of between 20 and 200 colonies.

INH serum concentrations After an overnight fast, INH was administered as crystalline powder (Fluka, Switzerland) with a glass of water; a light breakfast was allowed after the 2-h blood specimen was drawn. Blood for determination of INH concentrations was collected in ethylenediamine tetraacetic acid (EDTA)-coated tubes, immediately placed on ice, and delivered to the laboratory within 5 h. All samples were analysed on the same day that the specimen was received in the laboratory. INH concentrations were determined by high-performance liquid chromatography (HPLC) [33]. Analysis was by a Hewlett Packard 1090 L HPLC with an ultraviolet light detector, set at 340 nm, and an HP 3392 reporting integrator. Separation was on a Whatman Particil 5 C_8 250-mm column (i.d.=4.6 mm) maintained at 50°C with a flow rate of 1 ml/min. The mobile phase was composed of a mixture of 50 mM KH_2PO_4 (solvent A) in water and one volume part of isopropanol in four volume parts of acetonitrile (solvent B). The ratio of solvent A to solvent B was altered linearly over the 16-min time course of a run from 60:40 at the beginning to 30:70 at the end. Analysis of INH over the linear range of 0.5 µg/ml to 25 µg/ml (r= 0.99; n=10 per data point) showed a coefficient of variation of 1.98% and recovery from biological fluid of 100.01 ([standard deviation (SD) 1.13%]. For analysis of each batch of samples, a fresh calibration curve was set up with five freshly prepared standards within the range of 0.5–25 µg/ml. In no instance did standards deviate from the existing curve by more than 5%.

NAT2 genotyping of patients A 5-ml blood sample was collected from each patient into EDTA Vacutainer tubes for determination of the *NAT2* genotype profile [28]. Genomic DNA (gDNA) was extracted from blood by a salting-out procedure [22], and a 1,000 base pair (bp) sequence, coding for the NAT2 enzyme, was amplified by the polymerase chain reaction (PCR) technique. Each 100 µl PCR reaction volume contained 250 ng gDNA, 0.2 mM dNTP mixture, 2.5 mM $MgCl_2$, 0.3 µM primers, 1x PCR-buffer, and 1.25 U of Taq Polymerase I (Invitrogen, Carlsbad, USA). The 1,000-bp segment was analyzed by restriction fragment length polymorphism using six enzymes for appropriate cleavage (*Bam*H1, *Kpn*I, *Mp*I, *Taq*I, *Dde*I, and *Fok*I). The NAT2 primer sequences were used: Forward Primer: -5′-GACATTGAAGCATATTTTGAAAG-3′ Reverse Primer: -5′-ATGAAAGTATTTGATGTTTAGG-3′′. The *NAT2* DNA cleavage profiles yielded by each of the restriction enzymes were analyzed by gel electrophoresis using MetaPhor agarose (Biowhittaker, Rockland, MD, USA), and visualized using GelStar nucleic acid stain (Biowhittaker).

Statistical evaluation EBA was defined as EBA = $(Z_1 - Z_3)/2$, where Z_1 and Z_3 are logarithms of the CFU counts per milliliter sputum of the 16-h overnight sputum collections

Table 2 Influence of the *N*-acetyltransferase-2 (NAT2) subtype on the proportion of pulmonary tuberculosis patients receiving different doses of isoniazid (INH) and reaching an $AUC_{0-\infty}$ of ≥10.52 μg/ml per hour[a]

Dose mg/kg	NAT2 subtype								
	SS			FS			FF		
	n	$AUC_{0-\infty}$ ≥10.52 μg/ml	%	*n*	$AUC_{0-\infty}$ ≥10.52 μg/ml	%	*n*	$AUC_{0-\infty}$ ≥10.52 μg/ml	%
0.2	1	0	0	2	0	0	0	0	
1.5	8	2	25	7	1	14.3	1	0	0
3.0	4	4	100	6	2	33.3	3	1	33.3
5.0	21	21	100	27	26	96.3	12	4	33.3
6.0	6	6	100	5	5	100	10	9	90
10.0–12.0	21	21	100	29	29	100	16	16	100

SS homozygous slow NAT2, *FS* heterozygous fast NAT2, *FF* homozygous fast NAT2, *AUC* area under the curve

[a] Patients from study 1 [28] were evaluated following INH doses of both 5 mg/kg and 10 mg/kg body weight, and the number of patients in this table thus differs from that appearing in Table 1.

prior to the start of treatment (S1) and after 2 days of treatment (S3), respectively [19]. During the study of Parkin et al. [28], the pharmacokinetics of INH were studied in patients receiving 5 mg/kg INH and 10 mg/kg INH for the period 0–6 h after dosing. The two-compartment model as described by Bellman [2] was applied to these data and found to fit very well. This model was then employed to evaluate the data of the other patients for whom INH serum concentrations were available for only the period 2–5 h after dosing. A curve was fitted for each patient, and the $AUC_{0-\infty}$ calculated from the parameters of the fitted curve.

The relationship of EBA to INH $AUC_{0-\infty}$ and to the 2-h INH serum concentrations was then examined by exponential regression of EBA on INH $AUC_{0-\infty}$ and similarly by exponential regression of EBA on the 2-h INH serum concentration. The fitted regression curves were then used to estimate the INH $AUC_{0-\infty}$ and 2-h serum concentrations at which 90% of the asymptotic value of the fitted exponential regression EBA on $AUC_{0-\infty}$ or the 2-h serum concentration occurred [30]. In all regression analyses, the data from the relevant studies were pooled. The studies from which the data analyzed in this paper are drawn were all approved by the Institutional Review Board of the Faculty of Health Sciences of the University of Stellenbosch, and the patients all gave written informed consent for their participation in the studies.

Results

The mean maximum EBA estimated as the asymptote of the exponential regression is 0.569 (SE 0.047) and the EBA_{90} 0.512. The $AUC_{0-\infty}$ associated with the EBA_{90} is 10.52 (SE 3.69) μg/ml per hour. Similarly, the 2-h INH serum concentration associated with the EBA_{90} is 2.19 (SE 0.68) μg/ml. Table 2 summarizes the proportion of patients from each of the NAT2 subtypes receiving each of the INH doses that have an $AUC_{0-\infty}$ equal to or in excess of that associated with EBA_{90}. The results for all the dosage groups except 10–12 mg/kg are illustrated in Fig. 1. At the highest dose assessed (10–12 mg/kg), all patients had an INH $AUC_{0-\infty}$ in excess of 10.52 μg/ml per hour. At doses of 6 mg/kg and 5 mg/kg, however, one (10%) and eight (67%) of FF acetylators, respectively, failed to reach an AUC of 10.52 μg/ml per hour, as did one (4%) FS acetylator receiving a dose of 5 mg/kg.

Similarly, with regard to the 2-h INH serum concentrations, Table 3 summarizes the proportion of patients of each acetylator subtype and at each dose level that had a

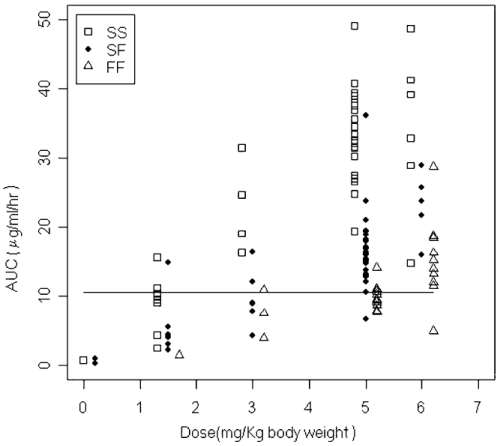


Fig. 1 The influence of the *N*-acetyltransferase-2 (NAT2) subtype on the $AUC_{0-\infty}$ of different doses of isoniazid (INH). The *horizontal line* indicates an $AUC_{0-\infty}$ of 10.52 μg/ml per hour, the threshold associated with an EBA_{90}. *AUC* area under the curve, *SS* homozygous slow NAT2, *FS* heterozygous fast NAT2, *FF* homozygous fast NAT2, *EBA* early bactericidal activity

2-h serum concentration ≥2.19 μg/ml, and the results for all of the dosage groups except the 10- to 12-mg/kg group, are illustrated in Fig. 2. At a dose of 10–12 mg/kg, all patients exceeded a serum concentration of 2.19 μg/ml, and at a dose of 6 mg/kg, all but two (20%) of the FF acetylators reached this concentration; however, at a dose of 5 mg/kg, nine of 12 (75%) FF actylators and 1 of 27 (4%) FS acetylators did not reach a concentration of 2.19 μg/ml.

Discussion

Following the introduction of INH into clinical use, a vigorous debate ensued as to the most appropriate dose of INH and the rhythm of its administration [1]. Early studies drew upon preliminary toxicity data to determine dose, and it was acknowledged that the question of the most appropriate dosage was "...still entirely unsettled," but that divided daily doses appeared to produce a greater systemic response than the administration of a single daily dose [34]. Early clinical studies in the UK followed a conservative path, giving a total daily dosage of INH of 200 mg in twice-daily 100-mg doses [37]. The debate was further complicated when it became clear that there were significant differences between individuals in their ability to eliminate INH [4, 17]. Subsequent studies showed that differences in the capacity of individuals to acetylate INH were responsible for the detected differences; individuals could be classified as fast or slow acetylators of INH, and there was also the suspicion that the fast group included a subgroup of very fast acetylators of INH [18]. Because of the considerable differences in the exposure of individuals to INH, there was immediate interest in the possible therapeutic consequences of these findings, and studies in the USA and UK addressed this issue [3, 5, 23, 26, 32].

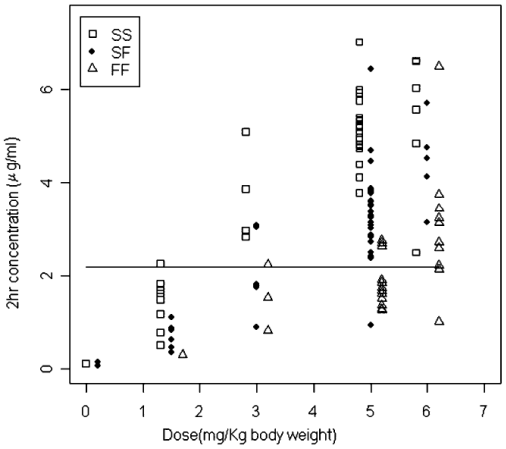


Fig. 2 The influence of the *N*-acetyltransferase-2 (NAT2) subtype on the 2-h isoniazid (INH) serum concentrations (μg/ml) reached in different dosage groups. The *horizontal line* indicates a 2-h serum concentration of 2.19 μg/ml, the threshold associated with an EBA_{90}. *SS* homozygous slow NAT2, *FS* heterozygous fast NAT2, *FF* homozygous fast NAT2, *EBA* early bactericidal activity

In the USA, several studies showed some advantage, particularly for the faster acetylators of INH, when higher doses were used, and a strong body of opinion advocated higher doses of INH for such individuals [23, 32]. This argument appeared to be finally settled when Harris reported a large study of a daily INH dose of 16 mg/kg given in thrice-daily divided doses that showed no particular advantage over a lower dose [15]. Between 1960 and 1966, the East African and British Medical Research Councils reported that increasing the dose of INH, accompanied by thiacetazone, from 200 mg daily in a single dose to 300 mg improved the response [10, 11], but

Table 3 Influence of the *N*-acetyltransferase-2 (NAT2) subtype on the proportion of pulmonary tuberculosis patients receiving different doses of isoniazid (INH) and reaching a 2-h serum concentration of ≥2.19 μg/ml[a]

Dose mg/kg	NAT2 subtype								
	SS			FS			FF		
	n	2 h ≥2.19 μg/ml	%	*n*	2 h ≥2.19 μg/ml	%	*n*	2 h ≥2.19 μg/ml	%
0.2	1	0	0.0	2	0	0.0	0	0	
1.5	8	1	12.5	7	0	0.0	1	0	0.0
3.0	4	4	100.0	6	2	33.3	3	1	33.3
5.0	21	21	100.0	27	26	96.3	12	3	25.0
6.0	6	6	100.0	5	5	100.0	10	8	80.0
10.0–12.0	21	21	100.0	29	29	100.0	16	16	100.0

SS homozygous slow NAT2, *FS* heterozygous fast NAT2, *FF* homozygous fast NAT2

[a] Patients from study 1 [28] were evaluated following INH doses of both 5 mg/kg and 10 mg/kg body weight, and the number of patients in this table thus differs from that appearing in Table 1.

that no better response was obtained by further increasing the dose of INH to 450 mg [12]. These last named studies did not, however, take cognizance of the NAT2 genotype or phenotype of the patients.

With regard to the killing of fully drug-sensitive bacilli, our own experience set out above supports this finding. Our EBA studies show a dose-related response, with a bactericidal effect of INH detectable from a total daily dose of 18 mg reaching a plateau at a total single daily dose of between 150 and 300 mg/day (3–6 mg/kg body weight for a patient weighting 50 kg) [6, 9]. Pharmacokinetic data from the studies analyzed in this paper estimate the INH serum concentrations and $AUC_{0-\infty}$ that are associated with 90% of the maximal killing (EBA_{90}) of metabolically active bacilli present in the sputum during the first 2 days of treatment. This occurs at a 2-h INH serum concentration of 2.19 μg/ml and an $AUC_{0-\infty}$ of 10.52 μg/ml per hour. At an INH dose of 6 mg/kg, even the great majority of FF acetylators of INH reach these targets. It is also evident that with any lowering of dose, there is a potential therapeutic price to be paid, in particular by the faster acetylators of INH. In the presence of full compliance with multidrug therapy, the demonstrated differences between FF and heterozygous fast acetylators and SS acetylators of INH are unlikely to be of therapeutic importance [13]. However, in the presence of noncompliance, whether overt or cryptic, [25] drug malabsorption [29], or widely spaced intermittent therapy [38, 39], the deficiencies of the faster NAT2 acetylators may well be exposed. In deciding the most appropriate dose of INH to recommend for clinical use, efficacy must be balanced with the possibility of toxicity. As regards efficacy our data set out above has, for the first time, provided objective criteria by which to measure the exposure of the patient to bactericidal concentrations of INH. It is evident from our data that SS acetylators of INH can be adequately treated with an INH dose of 3 mg/kg, whereas the FF acetylators require at least 6 mg/kg to reach the same degree of exposure to INH.

The major forms of toxicity associated with INH are hepatotoxicity and peripheral neuritis. Both of these are associated with slow acetylation of INH [16, 27, 36] and it has been proposed that determination of the NAT2 genotype should be used to assist in determining the dose of INH to be used in individual patients [20]. Thus, where the technology is available to determine the NAT2 genotype or phenotype, SS acetylators of INH could probably be safely treated with a daily INH dose of 3 mg/kg, thus ensuring a minimal probability of toxicity, but, none the less, reaching adequate serum INH concentrations to ensure optimum bactericidal activity, whereas faster acetylators of INH could receive 6 mg/kg, maximizing bactericidal activity without any overt danger of toxicity.

It must also be appreciated that INH will have its greatest use in tuberculosis control programs in the developing world. Here, large numbers of patients must be managed under program conditions with limited resources; an individualized approach to INH dosing is not feasible, and a compromise must be adopted. The dose of INH currently recommended by World Health Organization (WHO) of 5 mg/kg with a range of 4–6 mg/kg represents such a compromise [39]. In a population with a high incidence of rapid acetylators, it would be wise to use an INH dose of 6 mg/kg (300 mg in a 50-kg patient), whereas among a population of predominantly slow acetylators, 5 mg/kg might be more appropriate.

Finally, it must be acknowledged that our results were obtained in the Western Cape Province of South Africa among a population with a majority of faster acetylators of INH[30] and where a serious tuberculosis epidemic is being experienced, and these results may not be immediately applicable in other populations and other epidemiological situations. Also our findings apply to the EBA of INH and may not necessarily apply to the action of INH later in therapy; further exploration of our dosing proposals in the company of other antituberculosis agents and for the full course of treatment is desirable.

In conclusion our experience suggests that an INH dose of 3 mg/kg (150 mg for a 50-kg patient) will suffice to achieve the desired therapeutic objectives of antituberculosis treatment in SS acetylators of INH, but that FF acetylators require a dose of 6 mg/kg (300 mg for a 50-kg patient) to assure optimal bactericidal activity.

These studies were carried out in accordance with the requirements of the South African Medicines Control Council in force at the time that the studies were conducted.

Acknowledgement Dr. Don Parkin died during the preparation of the manuscript. His co-authors would like to acknowledge the significant role he played in all of our studies of the pharmacology of isoniazid.

References

1. Anonymous (1962) Isoniazid dosage. Tubercle 43:196–198
2. Bellman R (1983) Mathematical methods in medicine. Series in modern applied mathematics, vol 1. World Scientific, Singapore
3. Biehl JP (1957) Emergence of drug resistance as related to the dosage and metabolism of isoniazid. In: Transactions of 16th conference on the chemotherapy of tuberculosis. Veterans Administration Army Navy, Washington, DC, pp 108–113
4. Bönicke R, Reif W (1953) Enzymatische Inaktiviering von Isonikotinsaure hydrazide in menslichen und Tierschen Organismus. Arch Exp Pathol Pharmakol 220:321–333
5. Crofton J (1958) "Sputum conversion" and the metabolism of isoniazid. Am Rev Tuberc 77:869–871

6. Donald PR, Sirgel FA, Botha FJ et al (1997) The early bactericidal activity of isoniazid related to its dose size in pulmonary tuberculosis. Am J Respir Crit Care Med 156:895–900
7. Donald PR, Sirgel FA, Venter A et al (2001) Early bactericidal activity of amoxicillin in combination with clavulanic acid in patients with sputum smear-positive pulmonary tuberculosis. Scand J Infect Dis 33:466–469
8. Donald PR, Sirgel FA, Venter A et al (2002) The early bactericidal activity of streptomycin. Int J Tuberc Lung Dis 6:693–698
9. Donald PR, Sirgel FA, Venter A et al (2004) The influence of human N-acetyltransferase genotype on the early bactericidal activity of isoniazid. Clin Infect Dis 39:1425–1430
10. East African Hospitals/British Medical Research Council (1960) Comparative trial of isoniazid in combination with thiacetazone or a substituted diphenylthioures (SU1906) or PAS on the treatment of acute pulmonary tuberculosis in East Africa. Tubercle 41:399–423
11. East African/British Medical Research Council (1963) Isoniazid with thiacetazone in the treatment of pulmonary tuberculosis in East Africa. Second investigation. Tubercle 44:301–333
12. East African/British Medical Research Council (1966) Isoniazid with thiacetazone in the treatment of pulmonary tuberculosis in East Africa. Fourth investigation. Tubercle 47:315–339
13. Ellard G (1984) The potential clinical significance of isoniazid acetylator phenotype in the treatment of pulmonary tuberculosis. Tubercle 65:211–227
14. Elmendorf D, Cawthon WU, Muschenheim C, McDermott W (1952) The absorption, distribution, excretion and short-term toxicity of isonicotinic acid hydrazide (Nydrazid) in man. Am Rev Tuberc 429–442
15. Harris HW (1959) Report on study 16: high dosage INH. In: Transactions of the 18th conference on the chemotherapy of tuberculosis. Veterans Administration Army Navy, Washington, DC, pp 70–76
16. Hiratsuka M, Kishikawa Y, Takekuma Y et al (2002) Genotyping of the N-acetyltransferase 2 polymorphism in the prediction of adverse drug reactions to isoniazid in Japanese patients. Drug Metab Pharmacokinet 17:357–362
17. Hughes HB (1953) On the metabolic fate of isoniazid. J Pharmacol Exp Therap 109:444–454
18. Hughes HB, Biehl JP, Jones AP, Schmidt LH (1954) Metabolism of isoniazid in man as related to the occurrence of peripheral neuritis. Am Rev Tuberc 70:266–273
19. Jindani A, Aber VR, Edwards EA, Mitchison DA (1980) The early bactericidal activity of drugs in patients with pulmonary tuberculosis. Am Rev Respir Dis 121:939–949
20. Kinzig-Schippers M, Tomalik-Scharte D, Jetter A et al (2005) Should we use N-acetyltransferase type 2 genotyping to personalize isoniazid doses? Antimicrob Agents Chemother 49:1733–1738
21. Klee P (1952) Die Behandelung der Tuberkulose mit Neoteben (isonikotinsäuerhydrazid). Deutsche med Wchnschr 77:578–581
22. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acid Res 16:1215
23. Mitchell RS, Bell JC (1957) Clinical implications of isoniazid blood levels in pulmonary tuberculosis. New Eng J Med 257: 1066–1070
24. Mitchison DA (1979) Basic mechanisms of chemotherapy. Chest 76(Suppl):771–781
25. Mitchison DA (1998) How drug resistance emerges as result of poor compliance during short course chemotherapy for tuberculosis. Int J Tuberc Lung Dis 2:10–15
26. Mount FW, Ferebee SH (1954) United States Public Health Service cooperative investigation of antimicrobial therapy of tuberculosis. Report on thirty-two-week observations on combinations of isoniazid, streptomycin and para-aminosalicylic acid. Am Rev Tuberc 70:521–526
27. Ohno M, Yamaguchi I, Yamamoto I et al (2000) Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. Int J Tuberc Lung Dis 4:251–256
28. Parkin DP, Vandenplas S, Botha FJH et al (1997) Trimodality of isoniazid elimination. Phenotype and genotype in patients with tuberculosis. Am J Respir Crit Care Med 155:1717–1722
29. Peloquin CA, Nita AT, Burman WJ et al (1996) Low antituberculosis drug concentrations in patients with AIDS. Ann Pharmacother 30:919–925
30. R Development Core Team. R: 2006 A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org
31. Robitzek EH, Selikoff IJ (1952) Hydrazine derivatives of isonicotinic acid (Rimifon, Marsalid) in the treatment of active progressive caseous-pneumonic tuberculosis. Am Rev Tuberc 65:402–428
32. Russel WF, Kass I, Heaton AD, Dressler SH, Middlebrook G (1959) Combined drug treatment of tuberculosis. III Clinical application of the principles of appropriate and adequate chemotherapy to the treatment of pulmonary tuberculosis. J Clin Investigation 38:1366–1375
33. Seifart HI, Gent WI, Parkin DP, van Jaarsveld PP, Donald PR (1995) High-performance liquid chromatographic determination of isoniazid, acetylisoniazid and hydrazine in biological fluids. J Chromatogr [Br] 674:269–275
34. Selikoff IJ, Robitzek EH, Ornstein GG (1952) Treatment of pulmonary tuberculosis with hydrazine derivatives of isonicotinic acid. JAMA 150:973–980
35. Tuberculosis Chemotherapy Centre, Madras (1963) The prevention and treatment of isoniazid toxicity in the therapy of pulmonary tuberculosis. Bull Wld Hlth Org 38:455–475
36. Tuberculosis Chemotherapy Committee of the Medical Research Council (1953) Isoniazid in the treatment of pulmonary tuberculosis. Br Med J 1:521–536
37. Weiner M, Burman W, Vernon A et al (2003) Low isoniazid concentrations and outcome of tuberculosis treatment with once-weekly isoniazid and rifapentine. Am J Respir Crit Care Med 167:1341–1347
38. Weiner M, Benator D, Burman W et al (2005) Association between acquired rifamycin resistance and the pharmacokinetics of rifabutin and isoniazid among patients with HIV and tuberculosis. Clin Infect Dis 40:1481–1491
39. World Health Organization (2003) Standardized treatment regimen. In: Treatment of tuberculosis: guidelines for national programmes. WHO/CDS/TB/2003.313.Geneva. WHO.2003:37–38

CHAPTER 12

Conclusion

Drug metabolising enzymes (DMEs) are important determinants of individual-specific variations observed in the metabolism of xenobiotics and therapeutic drugs introduced into the human body. In this respect the study of Arylamine *N*-acetyltransferases (NAT1 and NAT2), is of particular importance, since these enzymes have the capacity to metabolise therapeutic drugs, such as *p*-amino-salicylate (PAS) [44], and isoniazid (INH) [18], respectively. Both of these compounds are important in the treatment of tuberculosis (TB) [18, 35, 39, 75]. Due to the increasing incidence of TB world-wide, the World Health Organisation (WHO) has declared TB a global health emergency (WHO Report 2006, and WHO TB Factsheet No.104) which clearly makes TB treatment a high priority. Effective TB treatment entails the elimination of the infecting organism, *Mycobacterium tuberculosis* (*Mtb*) from the human host [14], which in theory may be compromised by the activity of the NATs. Therefore it may be expedient to administer and monitor the anti-TB treatment in conjunction with individual's inherent acetylation capacity so as to maximise the treatment efficacy.

Previous studies have indicated that NAT2 enzyme activity can be correlated with specific single nucleotide polymorphisms (SNPs) in the gene [37, 161, 165, 168], and as such these SNPs can be used to rapidly screen for NAT2 acetylation activity in individuals. Similarly, numerous SNPs have been described for *NAT1* [76, 140], and the effect of these SNPs on NAT1 acetylation activity may be important for TB treatment since this enzyme acetylates PAS [44], which is used in the treatment of drug resistant TB. However, there is limited information available on *NAT1* alleles in African populations, and the effect of these *NAT1* SNPs has, to our knowledge, not been investigated in local South African communities.

Apart from the metabolism of compounds used for treating infectious diseases (such as INH for TB), the acetyltransferase enzymes also metabolise environmental compounds or xenobiotics. These xenobiotics or their metabolites may be risk factors for disease. In this context we investigated *NAT2* alleles as potential risk factors for the development of eye disease (cataract formation), as well as in a pilot study for high risk of cancer in our population. Clearly, the risk for the development of a chronic disease is unlikely to be linked to one single gene only, and thus in the cancer study we also investigated other genes as potential protection against xenobiotic damage.

If we wish to use the knowledge about the individual's acetylation status to decide on TB treatment, we would require a rapid acetylation assay to determine individual-specific acetylation. Currently, a relatively slow 7-SNP analysis method is widely used to determine the NAT2 acetylation status in individuals. We developed a rapid screening assay using the SNaPshot® technique to facilitate *NAT2* genotype analysis. Using this technique we were able to unambiguously analyse all 7 SNPs in a two tube multiplex assay. We were however unable to utilise a single tube multiplex protocol due to cross reactivity between the primers, which inhibited the analysis. The method realised a significant saving in time and labour, but requires more costly reagents, sophisticated laboratory equipment (DNA-sequencer) and technical expertise than the easier restriction enzyme based method (PCR-RFLP assay). Therefore in its current format, it may not be suitable for use in developing countries, where the need for the best treatment efficacy is greatest because of the high burden of TB disease.

Nevertheless, as a consequence of the Human Genome Project, DNA sequencing analysis is subject to ongoing optimisation, thereby becoming more affordable and user friendly [366, 367]. Currently sequencing costs are estimated at less than fifty American cents [368], and the technology is also able to achieve single molecule sequencing, using nanotechnology [369]. This may enable DNA sequencing centres to be set up in developing countries, which would facilitate the DNA sequencing analysis from remote and rural areas.

The SNaPshot® technique enables the high throughput screening of multiple sequence targets in the *NAT2* gene in multiple samples, which would improve the practical application of genotyping studies in pharmacogenetic studies. In addition, these results could be converted into a binary code for the purpose of SNP analysis (1 for the reference allele, and 0 for the polymorphic variant) which would facilitate the automation of the process. However, the methodology is very dependant upon primer sequence characteristics, and we were unable to adapt this technology to the genetic analysis of the *NAT1* SNPs, due to primer cross reactivity, as a result of the *NAT1* gene sequence characteristics (nucleotide order). In this respect, the technology may have a restricted application.

Genetic studies may often extrapolate findings in African American individuals to be applicable to Africans per se, despite the observation that the people on the African continent are the most genetically diverse group on earth [370, 371], and genetic results from one group would not necessarily be the same in other Africans. Unexpected phenotypes have been noted for different individuals and ethnic groups, and the source of these phenotypic variations have been linked to genetic polymorphisms [37, 48, 372, 373]. Furthermore, several reports support the hypothesis that disease susceptibility risk in diverse populations may be due to the relative proportions of certain alleles of DMEs [136, 374]. To test this hypothesis we investigated polymorphisms in two DMEs, namely Glutathione S-transferase (GST) T1, M1, P1 and Pi, which has been associated with cancer risk, and NAT2. Previous studies have established that the South African Black African (Xhosa) individuals have a high incidence of oesophageal cancer (OC) [375].

The analysis of GST polymorphism in the black and SAC groups returned interesting results. For example, the *GSTM1*0* (null) allele occurred at a frequency of 20% in both SAC and Xhosa groups, whilst the *GSTT1*0* allele was observed at a frequency of 41% and 57 % in the Xhosa and SAC groups respectively. The observed frequency for the *GSTM1*0* allele was similar in other African groups but statistically different in Caucasians and Chinese, where *GSTM1*0* occurred at a frequency of 40-60%. In the case of the *GSTT1*0* allele the frequency in Caucasian and Chinese was 10-24% and 64%, respectively. Moreover the combined *GSTM1*0*/*GSTT1*0* genotype was approximately 2-fold higher in Xhosa and SAC (at 11-13%) than in other African groups (at approximately 5%).

Similar observations were made for the *GSTP1*Ile^{105}/Val^{105} polymorphism, in which homozygous *GSTP1*Ile^{105}/Ile^{105} and *GSTP1*Va1^{105}l/Val^{105} occurred at a low frequency of 22 -28% respectively, whilst the heterozygous variant *GSTP1*Ile/Val occurred at a frequency of 51% in Xhosa individuals. By comparison, South African Venda had frequencies of 77%, 1% and 22% respectively.

The incidence of OC shows large disparities between Caucasians and Blacks, with a 2-fold greater survival rate in Caucasians than Blacks in North America [376]. OC is also prevalent amongst the Chinese, particularly in the north east of the country [377]. In

South Africa, the highest incidence of oesophageal cancer occurs in the Transkei, the cultural home land of the Xhosa [378]. These disparities in OC between ethnic groups has been an ongoing topic of research, and various genetic and environmental factors, such as dietary proclivities have been suggested to be risk factors in these ethnic groups.

In our investigation, we observed significant differences for specific *NAT* alleles between Caucasian, SAC and Xhosa ethnic groups. A similar finding was also evident in the frequency for GST alleles, in which differences are even evident amongst individuals belonging to the Venda and Xhosa Black African groups. Thus these results tend to support our idea of the inadmissibility of extrapolating the incidence of genetic polymorphisms in African populations. However, given that GST enzymes are involved in detoxification reactions, it may be that variations in *GST* alleles may play a role in increasing the risk oesophageal cancer in the Xhosa group.

Cataract formation and macular degeneration of the eye are major causes of visual impairment with which advancing age has been significantly associated. The aetiology of these conditions is multi-factorial and various "environmental" factors have been associated with cataract formation [379]. Drug metabolising enzymes such as NAT2, provide an important interface between environmental factors and the individual's inherent metabolism of chemical compounds introduced or made in the body, and these alleles were investigated to provide a better understanding of the aetiology of cataracts.

We were able to show that individuals belonging to the South African Coloured (SAC) group, harbouring the *NAT2* slow acetylator (S) allele, had a significantly higher risk for cataract formation [380]. Furthermore, this association was observed in both the homozygous slow (SS) and heterozygous intermediate (FS) genotypes, which represent the slow (SA) and intermediate (IA) acetylation phenotypes, respectively. NAT2 slow acetylation has also been indicated as a risk factor for cataract formation in a group of Turkish individuals [381]. Moreover, these authors also report that only the *NAT2*6A* allele was significantly associated with the condition in this population. Whilst a link

between cataract formation and NAT2 slow acetylation in particular is not intuitive, the NAT acetylation pathway is involved in the metabolism and subsequent detoxification of xenobiotic compounds containing amine and hydrazine moieties [46, 48], such as 4-aminobiphenyl, which is a constituent of tobacco smoke. In this context, it is interesting to note that smoking has been indicated as a risk factor for cataract formation [382]. However, it is not clear from the literature whether 4-aminobiphenyl is metabolised by NAT2 or NAT1 or both. This point may be moot, since it has been reported that both NAT1 and NAT2 genes are expressed in human lens epithelial cells [383]. Furthermore these authors indicate that the NAT genes can be inactivated by ultra-violet irradiation (UVB), which would further confirm the importance of the NAT enzymes in playing a protective role in the eye via the acetylation detoxification reactions.

The pharmacokinetic and pharmacodynamic characteristics of INH have been well established in studies of adult TB patients [384]. Similar data concerning children is generally lacking. In a high TB-incidence country like South Africa, an increasing number of children are affected with TB, and therefore a study was undertaken to facilitate a better understanding of INH metabolism in children. In a cohort of 64 children, with a median age of 3.8 years (range 1.8 – 7.8 years), we observed a trimodal model of INH elimination in children by both phenotype and genotype analyses [361]. These results are similar to earlier findings reported by our research group in adult individuals [63, 170]. More significantly, these recent results indicate a clear difference in the metabolism of INH between children and adults, as well as a faster elimination of INH in younger compared to older children. Furthermore, 23% were FA, 38% were IA, and SA was observed in 39% of the children. In this context it was also observed that in the FA group, 45% did not attain the recommended post-dose INH concentration of 1.5 mg/l, as required to facilitate treatment efficacy.

Given that the recommended treatment dosage range is based upon a measure of milligrams INH per kilogram body weight [301], it is believed that the faster INH elimination in children may be due to the greater proportion in liver mass with respect to

their body weight. Whilst this increased acetylation of INH in children has been shown in earlier phenotype studies, this study represents the first report to include the *NAT2* genotype analysis as confirmation of this earlier observation.

In the subsequent study of INH acetylation in children (chapter 10) we investigated the pharmacokinetics of INH in children given the dose of 4-6mg of INH per kg body weight as recommended by the World health Organization. In this latter we investigated the efficacy of this WHO recommended dosage, given the findings of the previous study (chapter 9) which indicated that children may have a greater INH elimination capacity when compared to adults. Our results indicate that children given this standard treatment regimen attained a peak concentration 58% lower than children given 8-10 mg/kg body weight. Furthermore, the majority of children (70%) on the standardised treatment regimen achieved a peak serum INH concentration of less than 3 mg/l which may be ineffective in the elimination of the *Mtb* organism and /or achieving treatment efficacy. This level of peak serum INH has been shown to be reached in adults given 300 mg of INH daily, which is effective in achieving the maximum cure rate. Hence it has been recommended that in children should be given a 8-10 mg/kg dose of INH in order to achieve peak serum concentrations comparable to that reached in adults. Furthermore, at the standard dose regimen infants may beat a greater risk of achieving o low peak serum concentration of INH than older children due to the increased rate of INH elimination reported by Schaaf *et al* [361].

The early bactericidal activity (EBA) of INH achieves a $\log_{10}$ reduction in the number of *Mtb* colony forming units during the first two days of treatment [173]. However, despite this characteristic, these authors showed that the individual-specific acetylation capacity has a significant influenced on the EBA of INH, and FA (bearing the homozygous RR

genotype) were deficient in achieving a INH peak serum concentration of 2-3 mg/ml to effect an optimal EBA. For this reason it has been advocated that FA should be treated with an increased INH dosage to off-set this apparent limitation, but that the most appropriate dose for treatment efficacy should be balanced against the well-known possibility of INH toxicity [18, 222, 385] due to the differing acetylation capacities inherent in FA, IA and SA. We therefore evaluated different doses of INH, ranging from 0.2 mg/kg to 12 mg/kg against the individual-specific NAT2 acetylation activity that would result in an optimal EBA in a group of SAC individuals.

The Pharmacokinetic results indicted that a 90% killing capacity (optimal EBA) occurred at an INH peak serum concentration of 2.19 µg/ml at 2 hours post-dose. This serum concentration was obtained in all individuals treated with an INH dose of 10-12 mg/kg body weight. At the 6 mg/kg dose 20% of FA failed to achieve this serum concentration, and at the lower dose of 5 mg/kg, 75% of FA did not exhibit the optimal EBA. On the other hand SA could be treated effectively with a 3 mg/kg INH dosage, with a minimal probability of drug toxicity. Similarly, no overt drug toxicity was observed among FA treated with INH at a dose of 6 mg/kg. Therefore in a population with a high incidence of FA such as the SAC, the 6 mg/kg dose would be appropriate, whereas in a population with predominantly SA, 5 mg/kg would be sufficient to achieve an optimal EBA, and hence would be more appropriate. This study presents data from a range of INH doses, and samples were only assayed over a short time period to assess the EBA, and hence no data on longer term outcomes were available However our results indicate that prior determination of the acetylation status in individuals would enable physicians to administer differing doses of INH in FA and SA to achieve the same optimal EBA.

In the TB case-control analysis of SAC individuals from a high TB-incidence area, the overall acetylation activity, as described by rapid (FA), intermediate (IA) and slow (SA) status, was similar in these groups. However, differences in *NAT2* allele (haplotype) frequencies were also observed between these groups. In the TB cases the *NAT2*4/2*12B, NAT2*12A/2*5B,* and *NAT2*4/2*6* IA genotypes were more prevalent

than in the TB-unaffected cohort. Here the common feature amongst these genotypes was the presence of the synonymous SNPs, 282C>T and 481C>T. In the case of the SA individuals the *NAT2*6A* haplotype was significantly associated with cases rather than with controls (table 10.6). However, the overall slow acetylation attributed by the *NAT2*6* class of alleles, did not differ between cases and controls, and hence it is unlikely that the *NAT2*6A* allele contributes to disease epidemiology. In this context, this observation remains without explanation, and may be merely a chance association. It is however curious that the same allele was also associated with cataract formation in SAC individuals.

We next extended our investigation of *NAT2* SNPs and their impact upon the acetylation status to SAC TB cases, treated with the standard regimen of first-line drugs, and intensively followed up. In contrast to the EBA study (chapter 11) the data for this study consisted of multiple samplings from individual patients, followed-up for two years after completion of the standard drug therapy. Our results show that the frequency distribution of the FA, IA and SA classes did not differ amongst responder (R) and non-responder (NR) individuals, as classified by their sputum smear microscopy (ZN-smear) and culture results after two months of the anti-TB treatment (standard directly observed treatment, sort course, (DOTs)). However, we observed a relatively low overall treatment conversion rate in our population, which can speculatively be ascribed to the relatively high proportion of rapid acetylators (homozygous (FA) and heterozygous (IA)) in this ethnic group compared to Caucasians, for example (table 1.6). Previous studies by our group have indicated that these rapid acetylators are unable to attain the optimum INH peak serum concentration required for maximum treatment efficacy, using the standardised DOTs dosing regimen of 5 mg of INH per kg body weight, because of the increased elimination of INH by this class of acetylators [173]. Hence the studies in chapters 4 and 11 are very different, and do not constitute follow-up studies as such. However, all information in science is interconnected and hence useful.

The NAT1 study represents the first analysis of *NAT1* genotypes in local ethnic groups to our knowledge. We therefore focussed our investigation upon the analysis of the 22 *NAT1* SNPs described in the NAT consensus database. Our initial results reflected a difference in the *NAT1* SNPs and alleles observed between the Caucasian and SAC cohorts, such that 16% of Caucasian and approximately 8% of SAC individuals had *NAT1* variants in exonic SNPs. The significance of this observation requires further investigation**,** particularly since many of these exonic variants, such as for example, SNP 350, 351G>C (rs 72554606), have yet to be characterised phenotypically, whilst others, such as the 777T>C (rs 4986991) SNP, encodes synonymous amino acids, which have been shown to have a definite effect on protein activity [305, 306, 327].

Interestingly, the 560G>A SNP, (which defines the loss of function allele *NAT1*14)*, was observed in both ethnic groups, albeit at a frequency of less than 1% in the SAC cohort and 6% in Caucasians. These rare variants are often overlooked or ignored in genetic analyses, because they are supposed to have no appreciable effect on the disease phenotype due to their very low frequency of occurrence. However, recent genetic studies, employing whole genome sequencing (GWAS), have shown that common (more frequently occurring) polymorphisms may only constitute a relatively small contribution to the disease phenotype [282-284, 386]. In the context of these results it would be premature to overlook the contribution of rare variants to genetic studies since they may define episodes of phenotypic variation which are observed clinically. It is these low frequency variants which may turn out to be better candidates to elucidate subtle phenotypic variations.

The most prevalent genotypes in the Caucasian and SAC cohorts were observed to be *NAT1*4/1*4*, *NAT1*4/1*10*, and *NAT1*1/1*10*, with the homozygous *NAT1*4/1*4* genotype predominating in Caucasians (54%), whilst the heterozygous *NAT1*4/1*10* genotype was the most prevalent in the SAC cohort (36%) (table 11.7). Furthermore, these genotype differences were statistically significant, indicating that their respective acetylation effects would make an important contribution to the overall NAT1 phenotype difference in these ethnic groups. By convention, the *NAT1*4* allele is the reference allele, encoding a rapid acetylation activity [339, 387], whilst the *NAT1*10* allele is presently considered to represent somewhat of an enigma. It has variously been described

as exhibiting an elevated acetylation activity in some studies, whilst in others it is assigned an acetylation status equivalent to that of the reference *NAT1*4* allele [75, 110, 159, 339, 340, 343, 387].

The *NAT1*10* allele (genotype) is defined by intronic SNPs (1088T>A and 1095C>A) that are located in the 3'-untranslated region (UTR) of the gene. Similarly, the 1095C>A SNP defines the *NAT1*3* allele. Whilst the functional effects of genetic polymorphisms in this domain of the gene are poorly understood [388], it has been shown that *NAT1* mRNA undergoes differential polyadenylation, initiated in the 3'-UTR of the gene [116]. It is further supposed that this may facilitate differential regulation of gene expression in different tissues [133-135]. Therefore, given the relatively high frequency of variants involving the 1088T>A and 1095C>A SNPs, as observed in both the Caucasian and SAC cohorts, and in the light of the reports just cited, these polymorphisms would constitute an important area for future research.

The investigation of the acetylation status in healthy individuals from the Caucasian, South African Coloured (SAC) and Black African (Xhosa) ethnic groups, from a tuberculosis (TB) endemic area, has shown that NAT2 slow acetylation (SA) predominates in Caucasian individuals, whilst intermediate acetylators (IA) were more prevalent in SAC and Xhosa individuals. The prevalence of IA in the SAC group may be particularly significant for TB treatment efficacy, since the outcome of the overall NAT2 acetylation activity in IA would be determined by the individual rapid and slow acetylator alleles inherent in the individual. Whilst the effects of individual *NAT2* slow alleles have been well characterised, the different *NAT2* rapid acetylator alleles may require further study, particularly with regard to the contribution (s) of synonymous SNPs present in the allele (haplotype). In view of the high burden of TB disease in SAC individuals [259], and because anti-TB treatment regimens are generally standardised based upon studies of Caucasian individuals, it is possible that the standard TB-treatment regimen would be less effective in SAC individuals, due to the increased acetylation (inactivation) of INH, by the relatively higher proportion of rapid acetylators in the latter population group,

compared to Caucasians. Furthermore, in light of the high burden of TB-disease in the SAC community, and the association reported between the *NAT2*6A* allele (haplotype) and drug induced hepatotoxicity in an Asian population [243], it may be important to closely monitor SAC TB-patients harbouring this genotype.

Our analysis of the *NAT* gene polymorphisms has highlighted the importance of individual SNPs and SNP-combinations (or haplotypes) in the acetylation activity of NAT2 in SAC individuals in particular, in South Africa. Similarly, the *NAT1* analysis has also reflected a prevalence of certain *NAT1* alleles in this ethnic group, compared to Caucasians. Whilst the functional effects of several of these *NAT* SNPs require further investigation, we have shown that the relative SNP and allele frequencies appear to be associated with certain ethnic groups, which could subsequently have an effect on the variation in acetylation activity in individuals from different ethnic groups, thereby affecting their response to certain therapeutic treatments, and/or their response to environmental insults.

Reference List

[1] Robitzek EH, Selikoff IJ. Hydrazine derivatives of isonicotinic acid (Rimifon, Marsalid) in the treatment of active progressive caseuos-pneumonic tuberculosis. *Am Rev Tuberc* 1952; **65:**402-428.

[2] Singh JA, Upshur R, Padayatchi N. XDR-TB in South Africa: No time for denial or complacency. *PLoS Medicine* 2007; **4:**e50.

[3] Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S *et al.* Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 2006; **103:**2869-2873.

[4] Menzies D. Tuberculosis crosses borders. *Int J Tuberc Lung Dis* 2000; **4:**S153-S159.

[5] Victor TC, de Haas PEW, Jordaan AM, van der Spuy GD, Richardson M, van Soolingen D *et al.* Molecular characteristics and global spread of *Mycobacterium tuberculosis* with a Western Cape F11 genotype. *J Clin Microbiol* 2004; **42:**769-772.

[6] Mokrousov I, Ly HM, Otten T, Lan NN, Vyshnevskyi B, Hoffner S *et al.* Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype: Clues from human phylogeography. *Genome Res* 2005; **15:**1357-1364.

[7] Hanekom M, van der Spuy GD, Gey van Pittius NC, McEvoy CRE, Ndabambi SL, Victor TC *et al.* Evidence that the spread of *Mycobacterium tuberculosis* strains with the Beijing genotype is human population dependant. *J Clin Microbiol* 2007; **45:**2263-2266.

[8] Salo WL, Aufderhiede AC, Buikstra J, Holcomb TA. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *PNAS* 1994; **91:**2091-2094.

[9] Zink AR, Sola C, Reischl U, Grabner W, Rastogi N, Wolf H *et al.* Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *Journal of Clinical Microbiology* 2003; **41:**359-367.

[10] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; **393:**537-544.

[11] Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends Microbiol* 2002; **10:**45-52.

[12] Dou HY, Tseng FC, Lu JJ, Jou R, Tsai SF, Chang JR *et al.* Associations of *Mycobacterium tuberculosis* genotypes with different ethnic and migratory populations in Taiwan. *Infection, Genetics and Evolution* 2008; **8:**323-330.

[13] Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infectious Diseases* 2007; **7:**328-337.

[14] Robitzek EH, Selikoff IJ, Ornstein GG. Chemotherpy of human tuberculosis with hydrazine derivatives of isonicotinic acid. *Q Bull Sea View Hosp* 1952; **31:**27-51.

[15] Zhang Y, Young DB. Molecular mechanisms of isoniazid: a drug at the front line of tuberculosis control. *Trends Microbiol* 1993; **1:**109-113.

[16] Mitchison DA. Role of individual drugs in the chemotherapy of tuberculosis. *Int J Tuberc Lung Dis* 2000; **4:**796-806.

[17] Hardman,JG, Limbird,LE, Molinoff,PB, Ruddon,RW, Gilman,AF. Goodman & Gilman's The Pharmacological Basis of Therapeutics.9th edition. NY, USA: McGraw-Hill; 1995.

[18] Hughes HB. On the metabolic fate of isoniazid. *J Pharmacol Exp Ther* 1953; **109:**444-452.

[19] Hughes HB, Schmidt LH, Biehl JP. The metabolism of isoniazid: Its implications in therapeutic use. *Conf Chemother Tuberc V A Armed Forces, 14th* 1955;217-221.

[20] Youatt J. Review of the action of isoniazid. *Am Rev Resp Dis* 1969; **99:**729-749.

[21] Banerjee A, Dubnau E, Quernard A, Balasubramanian V, Um KS, Wilson T *et al. InhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* 1994; **263:**227-230.

[22] Deretic V, Philipp W, Dhandayuthapani S, Mudd MH, Curcic R, Garbie T *et al.* Mycobacterium tuberculosis is a natural mutant with an inactivated oxidative-stress regulatory gene: implications for sensitivity to isoniazid. *Mol Microbiol* 1995; **17:**889-900.

[23] Kruger-Thiemer E. Isonicotinic acid: hypothesis of the antituberculous action of isoniazid. *Am Rev Tuberc* 1958; **77:**364-367.

[24] Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* 1992; **358:**591-593.

[25] Marrakchi H, Lanéelle G, Quémard A. InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. *Microbiology* 2000; **146:**289-296.

[26] Payton M, Auty R, Delgoda R, Everett M, Sim E. Cloning and characterization of arylamine N-acetyltransferase genes from Mycobacterium smegmatis and Mycobacterium tuberculosis: increased expression results in isoniazid resistance. *J Bacteriol* 1999; **181:**1343-1347.

[27] Payton M, Gifford C, Schartau P, Hagemeier C, Mushtaq A, Lucas S *et al.* Evidence towards the role of arylamine N-acetyltransferase in Mycobacterium smegmatis and development of a specific antiserum against the homologous enzyme of Mycobacterium tuberculosis. *Microbiology* 2001; **147:**3295-3302.

[28] Evans DA. N-acetyltransferase. *Pharmacology* 1989; **42:**157-234.

[29] Sim E, Payton M, Noble M, Minchin R. An update on genetic, structural and functional studies of arylamine N-acetyltransferases in eucaryotes and procaryotes. *Hum Mol Genet* 2000; **9:**2435-2441.

[30] Bhakta S, Besra GS, Upton AM, Parish T, Sholto-Douglas-Vernon C, Gibson KJ *et al.* Arylamine N-acetyltransferase is required for synthesis of mycolic acids and complex lipids in Mycobacterium bovis BCG and represents a novel drug target. *J Exp Med* 2004; **199:**1191-1199.

[31] Anderton MC, Bhakta S, Besra GS, Jeavons P, Eltis LD, Sim E. Characterisation of the putative operon containing arylamine N-acetyltransferase (nat) in *Mycobacterium bovis* BCG. *Mol Microbiol* 2006; **59:**181-192.

[32] Sinclair JC, Sandy J, Delgoda R, Sim E, Noble ME. Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat Struct Biol* 2000; **7:**560-564.

[33] Sandy J, Mushtaq A, Holton SJ, Schartau P, Noble MEM, Sim E. Investigation of the catalytic triad of arylamine *N*-acetyltransferase: essential residues required for acetyl transfer to arylamines. *Biochem J* 2005; **390:**115-123.

[34] Upton AM, Mushtaq A, Victor TC, Sampson SL, Sandy J, Smith DM *et al.* Arylamine N-acetyltransferase of Mycobacterium tuberculosis is a polymorphic enzyme and a site of isoniazid metabolism. *Mol Microbiol* 2001; **42:**309-317.

[35] Evans DA, White TA. Human acetylation polymorphism. *Journal of Laboratory and Clinical Medicine* 1964; **63:**394-402.

[36] Price Evans DA, Manley KA, McKusick VA. Genetic control of isoniazid metabolism in man. *Br Med J* 1960; **2:**485-491.

[37] Sunahara S, Urano M, Ogawa M. Genetic and geographic studies on isoniazid inactivation. *Science* 1961; **134:**1530-1531.

[38] Sunahara S, Urano M, Lin HT, Cheg TJ, Jarumilinda A. Further observations on trimodality of frequency distribution curve of biologically active isoniazid blood levels and "cline" infrequencies of alleles controlling isoniazid inactivation. *Acta Tuberc Pneumol Scand* 1963; **43:**181-195.

[39] Sunahara S, Urano M, Ogawa M, Yoshida S, Mukoyama H, Kawai R. Genetical aspect of isoniazid metabolism. *Jinrui Idengaku Zasshi* 1963; **187:**93-111.

[40] Deguchi T, Mashimo M, Suzuki T. Correlation between acetylator phenotypes and genotypes of polymorphic arylamine N-acetyltransferase in human liver. *J Biol Chem* 1990; **265:**12757-12760.

[41] Peters JH, Gordon GR, Brown P. The relationship between the capacities of human subjects to acetylate isoniazid, sulphanilamide and sulphamethazine. *Life Sci* 1965; **4:**99-107.

[42] Dreyfuss J, Bigger JT, Cohen AI, Schreiber EC. Metabolism of procainamide in rhesus monkey and man. *Clin Pharmacol Ther* 1972; **13:**366-371.

[43] Grant DM, Tang BK, Kalow WA. A simple test for acetylator phenotype using caffeine. *Br J Clin Pharmacol* 1984; **17:**459-464.

[44] Weber WW, Hein DW. N-acetylation pharmacogenetics. *Pharmacol Rev* 1985; **37:**25-79.

[45] Du Souich P, Erill S. Patterns of acetylation of procainamide and procainamide-derived p-aminobenzoic acid in man. *Eur J Clin Pharmacol* 1976; **10:**283-287.

[46] Weber,WW. The Acetylator genes and drug response.Oxford University Press, USA; 1987.

[47] Knight RA, Selin MJ, Harris HW. Genetic factors influencing isoniazid blood levels in humans. *Trans 18th Conf Chemother Tuberc Washington D C ,U S Veterans Adm Army navy:* 1959;52-58.

[48] Smith CA, Smith G, Wolf CR. Genetic polymorphisms in xenobiotic metabolism. *Eur J Cancer* 1994; **30A:**1921-1935.

[49] Minchin RF, Reeves PT, Teitel CH, McManus ME, Mojarrabi B, Ilett KF *et al. N*- and *O*-acetylation of aromatic and heterocyclic amine carcinogens by human monomorphic and polymorphic acetyltransferases expressed in COS-1 cells. *Biochem Biophys Res Commun* 1992; **185:**839-844.

[50] Meyer UA. Polymorphism of human acetyltransferases. *Environ Health Perspect* 1994; **102 Suppl 6:**213-216.

[51] Meyer UA, Zanger UM. Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annu Rev Pharmacol Toxicol* 1997; **37:**269-296.

[52] Watanabe M, Sofuni T, Nohmi T. Invovlvement of Cys^{69} residue in the catalytic mechanism of *N*-hydroxyarylamine *O*-acetyltransferase of *Salmonella typhimurium.* Sequence similarity at the amine acid level suggests a common catalytic mechanism of acetyltransferase for *S.typhimurium* and higher organisms. *J Biol Chem* 1992; **267:**8429-8436.

[53] Bessmann SF, Lipmann F. The enzymatic transacetylation between aromatic amines. *Arch Biochem Biophys Acta* 1953; **46:**252-254.

[54] Cleland WW. The kinetics of enzyme-catalysed reactions with two or more substrates or products. *Biochim Biophys Acta* 1963; **67:**104-137.

[55] Riddle B, Jencks WP. Acetyl-coenzyme A: arylamine N-acetyltransferase. Role of the acetyl-enzyme intermediate and the effects of substituents on the rate. *J Biol Chem* 1971; **246:**3250-3258.

[56] Andres HH, Vogel RS, Tarr GE, Johnson L, Weber WW. Purification, physicochemical, and kinetic properties of liver acetyl- CoA:arylamine N-acetyltransferase from rapid acetylator rabbits. *Mol Pharmacol* 1987; **31:**446-456.

[57] Andres HH, Kolb HJ, Schreiber RJ, Weiss L. Characterization of the active site, substrate specificity and kinetic properties of acetyl-CoA:arylamine N-acetyltransferase from pigeon liver. *Biochim Biophys Acta* 1983; **746:**193-201.

[58] Grant DM. Molecular genetics of the N-acetyltransferases. *Pharmacogenetics* 1993; **3:**45-50.

[59] Boteju LW, Hanna PE. Arylamine-nucleoside adduct formation: evidence for arylamine involvement in the reactions of an *N*-acetoxyarlyamine. *Chem Res Toxicol* 1994; **7:**684-689.

[60] Hein DW. Acetylator genotype and arylamine-induced carcinogenesis. *Biochim Biophys Acta* 1988; **948:**37-66.

[61] Guengerich FP, Gillam EM, Shimada T. New applications of bacterial systems to problems in toxicology. *Crit Rev Toxicol* 1996; **26:**551-583.

[62] Hickman D, Sim E. N-acetyltransferase polymorphism. Comparison of phenotype and genotype in humans. *Biochemical Pharmacology* 1991; **42:**1007-1014.

[63] Parkin DP, Vandenplas S, Botha FJ, Vandenplas ML, Seifart HI, van Helden PD *et al.* Trimodality of isoniazid elimination: phenotype and genotype in patients with tuberculosis. *Am J Respir Crit Care Med* 1997; **155:**1717-1722.

[64] Gross M, Kruisselbrink T, Anderson K, Lang N, McGovern P, Delongchamp R *et al.* Distribution and concordance of N-acetyltransferase genotype and phenotype in an American population. *Cancer Epidemiol Biomarkers Prev* 1999; **8:**683-692.

[65] Suzuki S, Muroishi Y, Nakanishi I, Oda Y. Relationship between genetic polymorphisms of drug-metabolizing enzymes (CYP1A1, CYP2E1, GSTM1, and NAT2), drinking habits, histological subtypes, and p53 gene point mutations in Japanese patients with gastric cancer. *J Gastroenterol* 2004; **39:**220-230.

[66] Grant DM, Lottspeich F, Meyer UA. Evidence for two closely related isozymes of arylamine N- acetyltransferase in human liver. *FEBS Lett* 1989; **244:**203-207.

[67] Meyer UA, Zanger UM, Skoda RC, Grant D, Blum M. Genetic polymorphisms of drug metabolism. *Prog Liver Dis* 1990; **9:**307-323.

[68] Hein DW, Doll MA, Fretland AJ, Leff MA, Webb SJ, Xiao GH *et al.* Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol Biomarkers Prev* 2000; **9:**29-42.

[69] Cribb AE, Grant DM, Miller MA, Spielberg SP. Expression of monomorphic arylamine N-acetyltransferase (NAT1) in human leukocytes. *J Pharmacol Exp Ther* 1991; **259:**1241-1246.

[70] Ohsako S, Deguchi T. Cloning and expression of cDNAs for polymorphic and monomorphic arylamine N-acetyltransferases from human liver. *J Biol Chem* 1990; **265:**4630-4634.

[71] Grant DM, Blum M, Beer M, Meyer UA. Monomorphic and polymorphic human arylamine N-acetyltransferases: a comparison of liver isozymes and expressed products of two cloned genes. *Mol Pharmacol* 1991; **39:**184-191.

[72] Weber WW, Vatsis KP. Individual variability in p-aminobenzoic acid N-acetylation by human N-acetyltransferase (NAT1) of peripheral blood. *Pharmacogenetics* 1993; **3:**209-212.

[73] Ward A, Hickman D, Gordon JW, Sim E. Arylamine N-acetyltransferase in human red blood cells. *Biochem Pharmacol* 1992; **44:**1099-1104.

[74] Butcher NJ, Ilett KF, Minchin RF. Functional polymorphism of the human arylamine N-acetyltransferase type 1 gene caused by C190T and G560A mutations. *Pharmacogenetics* 1998; **8:**67-72.

[75] Hughes NC, Janezic SA, McQueen KL, Jewett MA, Castranio T, Bell DA *et al.* Identification and characterization of variant alleles of human acetyltransferase NAT1 with defective function using p-aminosalicylate as an in-vivo and in-vitro probe. *Pharmacogenetics* 1998; **8:**55-66.

[76] Bruhn C, Brockmoller J, Cascorbi I, Roots I, Borchert HH. Correlation between genotype and phenotype of the human arylamine N-acetyltransferase type 1 (NAT1). *Biochem Pharmacol* 1999; **58:**1759-1764.

[77] Hickman D, Pope J, Patil SD, Fakis G, Smelt V, Stanley LA *et al.* Expression of arylamine N-acetyltransferase in human intestine. *Gut* 1998; **42:**402-409.

[78] Hein DW. N-Acetyltransferase genetics and their role in predisposition to aromatic and heterocyclic amine-induced carcinogenesis. *Toxicol Lett* 2000; **112-113:**349-356.

[79] Mahid SS, Colliver DW, Crawford NPS, Martini BD, Doll MA, Hein DW *et al.* Characterization of N-acetyltransferase 1 and 2 polymorphisms and haplotype analysis for inflammatory bowel disease and sporadic colorectal carcinoma. *BMC Med Genet* 2007; **8:**28.

[80] Hein DW. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mutat Res* 2002; **506-507:**65-77.

[81] Windmill KF, McKinnon RA, Zhu X, Gaedigk A, Grant DM, McManus ME. The role of xenobiotic metabolizing enzymes in arylamine toxicity and carcinogenesis: functional and localization studies. *Mutat Res* 1997; **376:**153-160.

[82] Sansarricq H. Recent advances and trends in leprosy research. *Experientia* 1977; **33:**114-117.

[83] Blum M, Grant DM, McBride W, Heim M, Meyer UA. Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression. *DNA Cell Biol* 1990; **9:**193-203.

[84] Hickman D, Risch A, Buckle V, Spurr NK, Jeremiah SJ, McCarthy A *et al.* Chromosomal localization of human genes for arylamine N- acetyltransferase. *Biochem J* 1994; **297 (Pt 3):**441-445.

[85] Franke S, Klawitz I, Schnakenberg E, Rommel B, Van d, V, Bullerdiek J *et al.* Isolation and mapping of a cosmid clone containing the human *NAT2* gene. *Biochem Biophys Res Commun* 1994; **199:**52-55.

[86] Matas N, Thygesen P, Stacey M, Risch A, Sim E. Mapping AAC1, AAC2 and AACP, the genes for arylamine N-acetyltransferases, carcinogen metabolising enzymes on human chromosome 8p22, a region frequently deleted in tumours. *Cytogenet Cell Genet* 1997; **77:**290-295.

[87] Ebisawa T, Deguchi T. Structure and restriction fragment length polymorphism of genes for human liver arylamine N-acetyltransferases. *Biochem Biophys Res Commun* 1991; **177:**1252-1257.

[88] Dupret JM, Goodfellow GH, Janezic SA, Grant DM. Structure-function studies of human arylamine N-acetyltransferases NAT1 and NAT2. Functional analysis of recombinant NAT1/NAT2 chimeras expressed in Escherichia coli. *J Biol Chem* 1994; **269:**26830-26835.

[89] Stanley LA, Coroneos E, Cuff R, Hickman D, Ward A, Sim E. Immunochemical detection of arylamine N-acetyltransferase in normal and neoplastic bladder. *J Histochem Cytochem* 1996; **44:**1059-1067.

[90] Risch A, Smelt V, Lane D, Stanley L, van der SW, Ward A *et al.* Arylamine N-acetyltransferase in erythrocytes of cystic fibrosis patients. *Pharmacol Toxicol* 1996; **78:**235-240.

[91] Smelt VA, Mardon HJ, Redman CW, Sim E. Acetylation of arylamines by the placenta. *Eur J Drug Metab Pharmacokinet* 1997; **22:**403-408.

[92] Pacifici GM, Bencini C, Rane A. Acetyltransferases in humans: development and tissue distribution. *Pharmacology* 1986; **32:**283-291.

[93] Jenne JW. Partial purification and properties of the isoniazid transacetylase in human liver. Its relationship to the acetylation of *p*-aminosalicyclic acid. *J Clin Invest* 1965; **44:**1992-2005.

[94] Ilett KF, Ingram DM, Carpenter DS, Teitel CH, Lang NP, Kadlubar FF *et al.* Expression of monomorphic and polymorphic *N*-acetyltransferases in human colon. *Biochem Pharmacol* 1994; **47:**914-917.

[95] Grant DM, Blum M, Demierre A, Meyer UA. Nucleotide sequence of an intronless gene for a human arylamine N- acetyltransferase related to polymorphic drug acetylation. *Nucleic Acids Res* 1989; **17:**3978.

[96] Grant DM, Goodfellow GH, Sugamori K, Durette K. Pharmacogenetics of the human arylamine N-acetyltransferases. *Pharmacology* 2000; **61:**204-211.

[97] Kawamura A, Graham J, Mushtaq A, Tsiftsoglou SA, Vath GM, Hanna PE *et al.* Eukaryotic arylamine N-acetyltransferase. Investigation of substrate specificity by high-throughput screening. *Biochem Pharmacol* 2005; **69:**347-359.

[98] Ohsako S, Ohtomi M, Sakamoto Y, Uyemura K, Deguchi T. Arylamine N-acetyltransferase from chicken liver II. Cloning of cDNA and expression in Chinese hamster ovary cells. *J Biol Chem* 1988; **263:**7534-7538.

[99] Frymoyer JW, Jacox RF. Studies of genetically controlled sulfadiazine acetylation in rabbit livers: Possible identification of the heterozygous trait. *J Lab Clin Med* 1963; **62:**905-909.

[100] Blum M, Grant DM, Demierre A, Meyer UA. Nucleotide sequence of a full-length cDNA for arylamine N- acetyltransferase from rabbit liver. *Nucleic Acids Res* 1989; **17:**3589.

[101] Doll MA, Hein DW. Cloning, sequencing and expression of NAT1 and NAT2 encoding genes from rapid and slow acetylator inbred rats. *Pharmacogenetics* 1995; **5:**247-251.

[102] Trepanier LA, Cribb AE, Spielberg SP, Ray K. Deficiency of cytosolic arylamine N-acetylation in the domestic cat and wild felids caused by the presence of a single NAT1-like gene. *Pharmacogenetics* 1998; **8:**169-179.

[103] Sinclair JC, Delgoda R, Noble ME, Jarmin S, Goh NK, Sim E. Purification, characterisation, and crystallization of an *N*-hydroxyarylamine *O*-acetyltransferase from *Salmonella typhimurium*. *Protein Expr Purif* 1998; **12:**371-380.

[104] Trepanier LA, Ray K, Winand NJ, Spielberg SP, Cribb AE. Cytosolic arylamine N-acetyltransferase (NAT) deficiency in the dog and other canids due to an absence of NAT genes. *Biochem Pharmacol* 1997; **54:**73-80.

[105] Dupret JM, Grant DM. Site-directed mutagenesis of recombinant human arylamine N- acetyltransferase expressed in Escherichia coli. Evidence for direct involvement of Cys68 in the catalytic mechanism of polymorphic human NAT2. *J Biol Chem* 1992; **267:**7381-7385.

[106] Rodrigues-Lima F, Delomenie C, Goodfellow GH, Grant DM, Dupret JM. Homology modelling and structural analysis of human arylamine N-

acetyltransferase NAT1: evidence for the conservation of a cysteine protease catalytic domain and an active-site loop. *Biochem J* 2001; **356:**327-334.

[107] Wu H, Dombrovsky L, Tempel W, Martin F, Loppnau P, Goodfellow GH *et al.* Structural basis of substrate-binding specificity of human arylamine *N*-acetyltransferases. *J Biol Chem* 2007; **282:**30189-30197.

[108] Walraven JM, Trent JO, Hein DW. Computational and experimental analyses of mammalian arylamine N-acetyltransferase structure and function. *Drug Metab Dispos* 2007; **35:**1001-1007.

[109] Sim E, Walters K, Boukouvala S. Arylamine N-acetyltransferases: from structure to function. *Drug Metab Rev* 2008; **40:**479-510.

[110] Payton MA, Sim E. Genotyping human arylamine N-acetyltransferase type 1 (NAT1): the identification of two novel allelic variants. *Biochem Pharmacol* 1998; **55:**361-366.

[111] Goodfellow GH, Dupret JM, Grant DM. Identification of amino acids imparting acceptor substrate selectivity to human arylamine acetyltransferases NAT1 and NAT2. *Biochem J* 2000; **348 Pt 1:**159-166.

[112] Butcher NJ, Arulpragasam A, Minchin RF. Proteasomal degradation of N-acetyltransferase 1 is prevented by acetylation of the active site cysteine: a mechanism for the slow acetylator phenotype and substrate-dependent down-regulation. *J Biol Chem* 2004; **279:**22131-22137.

[113] Cornish VA, Pinter K, Boukouvala S, Johnson N, Labrousse C, Payton M *et al.* Generation and analysis of mice with a targeted disruption of the arylamine *N*-acetyltransferase type 2 gene. *Pharmacogenomics J* 2003; **3:**169-177.

[114] Boukouvala S, Fakis G. Arylamine N-acetyltransferases: what we learn from genes and genomes. *Drug Metab Rev* 2005; **37:**511-564.

[115] Fakis G, Boukouvala S, Buckle V, Payton M, Denning C, Sim E. Chromosome mapping of the genes for murine arylamine N-acetyltransferases (NATs), enzymes involved in the metabolism of carcinogens: identification of a novel upstream noncoding exon for murine Nat2. *Cytogenet Cell Genet* 2000; **90:**134-138.

[116] Boukouvala S, Sim E. Structural analysis of the genes for human arylamine N-acetyltransferases and characterisation of alternative transcripts. *Basic Clin Pharmacol Toxicol* 2005; **96:**343-351.

[117] Boukouvala S, Price N, Sim E. Identification and functional characterization of novel polymorphisms associated with the genes for arylamine N-acetyltransferases in mice. *Pharmacogenetics* 2002; **12:**385-394.

[118] Loehle JA, Cornish V, Wakefield L, Doll MA, Neale JR, Zang Y *et al.* N-acetyltransferase (Nat) 1 and 2 expression in Nat2 knockout mice. *J Pharmacol Exp Ther* 2006; **319:**724-728.

[119] Windmill KF, Gaedigk A, Hall PM, Samaratunga H, Grant DM, McManus ME. Localization of N-acetyltransferases NAT1 and NAT2 in human tissues. *Toxicol Sci* 2000; **54:**19-29.

[120] Husain A, Barker DF, States JC, Doll MA, Hein DW. Identification of the major promoter and non-coding exons of the human arylamine N-acetyltransferase 1 gene (*NAT1*). *Pharmacogenetics* 2004; **14:**397-406.

[121] Barker DF, Husain A, Neale JR, Martini BD, Zhang X, Doll MA *et al.* Functional properties of an alternative, tissue-specific promoter for human arylamine *N*-acetyltransferase 1. *Pharmacogenetics* 2006; **16:**525.

[122] Husain A, Zhang X, Doll MA, States JC, Barker DF, Hein DW. Functional analysis of the human *N-Acetytransferase 1* major promoter: Quantitation of tissue expression and identification of critical sequence elements. *Drug Metab Dispos* 2007; **35:**1649-1656.

[123] Butcher NJ, Arulpragasam A, Pope C, Minchin RF. Identification of a minimal promoter sequence for the human N-acetyltransferase Type I gene that binds AP-1 (activator protein 1) and YY-1 (Yin and Yang 1). *Biochem J* 2003; **376:**441-448.

[124] Philipsen S, Suske G. A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res* 1999; **27:**2991-3000.

[125] Mitchell KR, Warshawsky D. Xenobiotic inducible regions of the human arylamine N-acetyltransferase 1 and 2 genes. *Toxicol Lett* 2003; **139:**11-23.

[126] Minchin RF, Hanna PE, Dupret JM, Wagner CR, Rodrigues-Lima F, Butcher NJ. Arylamine *N*-acetyltransferase 1. *The International Journal of Biochemistry and Cell Biology* 2007; **39:**1999-2005.

[127] Husain A, Zhang X, Doll MA, States JC, Barker DF, Hein DW. Identification of *N*-acetyltransferase 2 (*NAT2*) transcription start sites and quantitation of *NAT2*-specific mRNA in human tissues. *Drug Metab Dispos* 2007; **35:**721-727.

[128] Sim E, Lack N, Wang CJ, Long H, Westwood I, Fullam E *et al.* Arylamine N-acetyltransferases: structural and functional implications of polymorphisms. *Toxicology* 2008; **254:**170-183.

[129] Yuliwulandari R, Sachrowardi Q, Nishida N, Takasu M, Batubara L, Susmiarsih TP *et al.* Polymorphisms of promoter and coding regions of arylamine *N*-acetyltransferase 2 (*NAT2*) gene in the Indonesian population: proposal for a new nomenclature. *J Hum Genet* 2008; **53:**201-209.

[130] Land SJ, Jones RF, King CM. Biochemical and genetic analysis of two acetyltransferases from hamster tissues that can metabolize aromatic amine derivatives. *Carcinogenesis* 1994; **15:**1585-1595.

[131] Ebisawa T, Sasaki Y, Deguchi T. Complementary DNAs for two arylamine N-acetyltransferases with identical 5' non-coding regions from rat pineal gland. *Eur J Biochem* 1995; **228:**129-137.

[132] Sasaki Y, Ohsako S, Deguchi T. Molecular and genetic analyses of arylamine N-acetyltransferase polymorphism of rabbit liver. *J Biol Chem* 1991; **266:**13243-13250.

[133] Smith CWJ, Patton JG, Nadal-Ginard B. Alternative splicing in the control of gene expression. *Annu Rev Genet* 1989; **23:**527-577.

[134] Sonenberg N. mRNA translation: influence of the 5' and 3' untranslated regions. *Curr Opin Genet Dev* 1994; **4:**310-315.

[135] Rampazzo A, Pivotto F, Occhi G, Tiso N, Bortoluzzi S, Rowen L *et al.* Characterization of C14orf4, a novel intronless human gene containing a polyglutamine repeat, mapped to the ARVD1 critical region. *Biochem Biophys Res Commun* 2000; **278:**766-774.

[136] Risch A, Wallace DM, Bathers S, Sim E. Slow N-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer. *Hum Mol Genet* 1995; **4:**231-236.

[137] Zhu J, Chang P, Bondy ML, Sahin AA, Singletary SE, Takahashi S *et al.* Detection of 2-amino-1-methyl-6-phenylimidazo[4,5-b]-pyridine-DNA adducts in normal breast tissues and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2003; **12:**830-837.

[138] Le Marchand L, Hankin JH, Wilkens LR, Pierce LM, Franke A, Kolonel LN *et al.* Combined effects of well-done red meat, smoking, and rapid *N*-acetyltransferase 2 and CYP1A2 phenotypes in increasing colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 2001; **10:**1259-1266.

[139] Costa S, Pinto D, Morais A, Vasconcelos A, Oliveira J, Lopes C *et al.* Acetylation genotype and the genetic susceptibility to prostate cancer in a southern European population. *Prostate* 2005; **64:**246-252.

[140] Hein DW. *N*-acetyltransferase SNPs: emerging concepts serve as a paradigm for understanding complexities of personalized medicine. *Expert Opin Drug Metab Toxicol* 2009; **5:**353-366.

[141] Bell DA, Badawi AF, Lang NP, Ilett KF, Kadlubar FF, Hirvonen A. Polymorphism in the N-acetyltransferase 1 (NAT1) polyadenylation signal: association of NAT1*10 allele with higher N-acetylation activity in bladder and colon tissue. *Cancer Res* 1995; **55:**5226-5229.

[142] Probst-Hensch NM, Haile RW, Li DS, Sakamoto GT, Louie AD, Lin BK *et al.* Lack of association between the polyadenylation polymorphism in the NAT1 (acetyltransferase 1) gene and colorectal adenomas. *Carcinogenesis* 1996; **17:**2125-2129.

[143] Bouchardy C, Mitrunen K, Wikman H, Husgafvel-Pursiainen K, Dayer P, Benhamou S *et al.* N-acetyltransferase NAT1 and NAT2 genotypes and lung cancer risk. *Pharmacogenetics* 1998; **8:**291-298.

[144] Lo-Guidice JM, Allorge D, Chevalier D, Debuysere H, Fazio F, Lafitte LJ *et al.* Molecular analysis of the N-acetyltransferase 1 gene (NAT1*) using polymerase chain reaction-restriction fragment-single strand conformation polymorphism assay. *Pharmacogenetics* 2000; **10:**293-300.

[145] Karolyi J, Erickson RJ, Liu S, Killewald L. Major effects on teratogen-induced facial clefting in mice determined by a single genetic region. *Genetics* 1990; **126:**201-205.

[146] Wakefield L, Long H, Lack N, Sim E. Ocular defects associated with a null mutation in the mouse arylamine N-acetyltransferase 2 gene. *Mamm Genome* 2007; **18:**270-276.

[147] Cascorbi I, Roots I, Brockmoller J. Association of NAT1 and NAT2 polymorphisms to urinary bladder cancer: significantly reduced risk in subjects with NAT1*10. *Cancer Res* 2001; **61:**5051-5056.

[148] Hein DW, Leff MA, Ishibe N, Sinha R, Frazier HA, Doll MA *et al.* Association of prostate cancer with rapid N-acetyltransferase 1 (NAT1*10) in combination with slow N-acetyltransferase 2 acetylator genotypes in a pilot case-control study. *Environ Mol Mutagen* 2002; **40:**161-167.

[149] Boissy RJ, Watson MA, Umbach DM, Deakin M, Elder J, Strange RC *et al.* A pilot study investigating the role of NAT1 and NAT2 polymorphisms in gastric adenocarcinoma. *Int J Cancer* 2000; **87:**507-511.

[150] Butcher NJ, Ilett KF, Minchin RF. Substrate-dependent regulation of human arylamine N-acetyltransferase-1 in cultured cells. *Mol Pharmacol* 2000; **57:**468-473.

[151] Rodrigues-Lima F, Dupret JM. Regulation of the activity of the human drug metabolizing enzyme arylamine N-acetyltransferase 1: role of genetic and non genetic factors. *Curr Pharm Des* 2004; **10:**2519-2524.

[152] Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001; **98:**10869-10874.

[153] Adam PJ, Berry J, Loader JA, Tyson KL, Craggs G, Smith P *et al.* Arylamine N-acetyltransferase-1 is highly expressed in breast cancers and conveys enhanced growth and resistance to etoposide in vitro. *Mol Cancer Res* 2003; **1:**826-835.

[154] Zhao B, Lee EJ, Yeoh PN, Gong NH. Detection of mutations and polymorphism of N-acetyltransferase 1 gene in Indian, Malay and Chinese populations. *Pharmacogenetics* 1998; **8:**299-304.

[155] Loktionov A, Moore W, Spencer SP, Vorster H, Nell T, O'Neill IK *et al.* Differences in N-acetylation genotypes between Caucasians and Black South Africans: implications for cancer prevention. *Cancer Detect Prev* 2002; **26:**15-22.

[156] Dhaini HR, Levy GN. Arylamine N-acetyltransferase 1 (NAT1) genotypes in a Lebanese population. *Pharmacogenetics* 2000; **10:**79-83.

[157] Kukongviriyapan V, Prawan A, Warasiha B, Tassaneyakul W, Aiemsa-Ard J. Polymorphism of N-acetyltransferase 1 and correlation between genotype and phenotype in a Thai population. *Eur J Clin Pharmacol* 2003; **59:**277-281.

[158] Zhangwei X, Jianming X, Qiao M, Xinhua X. *N*-acetyltransferase-1 gene polymorphisms and correlation between genotype and its activity in a central Chinese Han population. *Clinica Chimica Acta* 2006; **371:**85-91.

[159] Yang M, Katoh T, Delongchamp R, Ozawa S, Kohshi K, Kawamoto T. Relationship between NAT1 genotype and phenotype in a Japanese population. *Pharmacogenetics* 2000; **10:**225-232.

[160] Hsieh FI, Pu YS, Chern HD, Hsu LI, Chiou HY, Chen CJ. Genetic polymorphisms of N-acetyltransferase 1 and 2 and risk of cigarette smoking-related bladder cancer. *Br J Cancer* 1999; **81:**537-541.

[161] Sekine A, Saito S, Iida A, Mitsunobu Y, Higuchi S, Harigae S *et al.* Identification of single-nucleotide polymorphisms (SNPs) of human N-acetyltransferase genes NAT1, NAT2, AANAT, ARD1 and L1CAM in the Japanese population. *J Hum Genet* 2001; **46:**314-319.

[162] Biehl JP. Emergence of drug resistance as related to the dosage and metabolism of isoniazid. *Trans 16th Conf Chemother Tuberc Washington D C , U S Veterans Adm Army Navy* 1957;108-113.

[163] Mashimo M, Suzuki T, Abe M, Deguchi T. Molecular genotyping of N-acetylation polymorphism to predict phenotype. *Hum Genet* 1992; **90:**139-143.

[164] Bell DA, Taylor JA, Butler MA, Stephens EA, Wiest J, Brubaker LH *et al.* Genotype/phenotype discordance for human arylamine N-acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African-Americans. *Carcinogenesis* 1993; **14:**1689-1692.

[165] Lin HJ, Han CY, Lin BK, Hardy S. Ethnic distribution of slow acetylator mutations in the polymorphic N- acetyltransferase (NAT2) gene. *Pharmacogenetics* 1994; **4:**125-134.

[166] Straka RJ, Hansen SR, Benson SR, Walker PF. Predominance of slow acetylators of N-acetyltransferase in a Hmong population residing in the United States. *J Clin Pharmacol* 1996; **36:**740-747.

[167] Patin E, Harmant C, Kidd KK, Kidd J, Froment A, Mehdi SQ *et al.* Sub-Saharan African coding sequence variation and haplotype diversity at the *NAT2* gene. *Hum Mutat* 2006; **27:**720-730.

[168] Fretland AJ, Leff MA, Doll MA, Hein DW. Functional characterization of human N-acetyltransferase 2 (NAT2) single nucleotide polymorphisms. *Pharmacogenetics* 2001; **11:**207-215.

[169] Hickman D. Genotyping human polymorphic arylamine N-acetyltransferase: identification of new slow allotypic variants. *Pharmacogenetics* 1992; **2:**217-226.

[170] Parkin DP, Vandenplas ML, Donald PR, Vandenplas S, Botha FJ, Seifart HI *et al.* Distribution of acetylator genotypes in the coloured population of the Western Cape region of South Africa. *South African Journal of Science* 1998; **94:**404-406.

[171] Hamdy SI, Hiratsuka M, Narahara K, Endo N, El-Enany M, Moursi N *et al.* Genotype and allele frequencies of TPMT, NAT2, GST, SULT1A1 and MDR-1 in the Egyptian population. *Br J Clin Pharmacol* 2003; **55:**560-569.

[172] Anitha A, Banerjee M. Arylamine N-acetyltransferase 2 polymorphism in the ethnic populations of South India. *Int J Mol Med* 2003; **11:**125-131.

[173] Donald PR, Sirgel FA, Venter A, Parkin DP, Seifart HI, Van de Wal BW *et al.* The influence of human N-acetyltransferase genotype on the early bactericidal activity of isoniazid. *Clin Infect Dis* 2004; **39:**1425-1430.

[174] Cascorbi I, Drakoulis N, Brockmoller J, Maurer A, Sperling K, Roots I. Arylamine N-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity. *Am J Hum Genet* 1995; **57:**581-592.

[175] Hein DW, Ferguson RJ, Doll MA, Rustan TD, Gray K. Molecular genetics of human polymorphic N-acetyltransferase: enzymatic analysis of 15 recombinant wild-type, mutant, and chimeric NAT2 allozymes. *Hum Mol Genet* 1994; **3:**729-734.

[176] Zang Y, Zhao S, Doll MA, States JC, Hein DW. The T341C (Ile114Thr) polymorphism of N-acetyltransferase 2 yields slow acetylator phenotype by enhanced protein degradation. *Pharmacogenetics* 2004; **14:**717-723.

[177] Zang Y, Doll MA, Zhao S, States JC, Hein DW. Functional characterization of single-nucleotide polymorphisms and haplotypes of human N-acetyltransferase 2. *Carcinogenesis* 2007; **28:**1665-1671.

[178] Hickman D, Palamanda JR, Unadkat JD, Sim E. Enzyme kinetic properties of human recombinant arylamine N- acetyltransferase 2 allotypic variants expressed in Escherichia coli. *Biochem Pharmacol* 1995; **50:**697-703.

[179] Zhu Y, Doll MA, Hein DW. Functional genomics of $C^{190}T$ single nucleotide polymorphism in human *N*-acetyltransferase 2. *Biol Chem* 2002; **383:**983-987.

[180] Zang Y, Zhao S, Doll MA, Christopher State, Hein DW. Functional characterization of the A411T (L137F) and G364A (D122N) genetic polymorphisms in human N-acetyltransferase 2. *Pharmacogenet Genomics* 2007; **17:**37-45.

[181] Hein DW, Doll MA, Rustan TD, Ferguson RJ. Metabolic activation of N-hydroxyarylamines and N-hydroxyarylamides by 16 recombinant human NAT2 allozymes: effects of 7 specific NAT2 nucleic acid substitutions. *Cancer Res* 1995; **55:**3531-3536.

[182] Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A *et al.* Human acetyltransferase polymorphisms. *Mutat Res* 1997; **376:**61-70.

[183] Shiokawa S, Yasuda M, Nobunaga M. Genotypes of polymorphic arylamine N-acetyltransferase in systemic lupus erythematosus. *Arthritis Rheum* 1992; **35:**1397-1399.

[184] Bozok C, V, Erer OF, Kosova B, Ozdemir I, Topcuoglu N, Aktogu S *et al.* Determining the relation between N-acetyltransferase-2 acetylator phenotype and antituberculosis drug induced hepatitis by molecular biologic tests. *Tuberk Toraks* 2008; **56:**81-86.

[185] Lin HJ, Han CY, Lin BK, Hardy S. Slow acetylator mutations in the human polymorphic N-acetyltransferase gene in 786 Asians, blacks, Hispanics, and whites: application to metabolic epidemiology. *Am J Hum Genet* 1993; **52:**827-834.

[186] Fuselli S, Gilman RH, Chanock SJ, Bonatto SL, de Stefano G, Evans CA *et al.* Analysis of nucleotide diversity of NAT2 coding region reveals homogeneity across Native American populations and high intra-population diversity. *The Pharmacogenomics Journal* 2007; **7:**144-152.

[187] Rabstein S, Unfried K, Ranft U, Illig T, Kolz M, Rihs HP *et al.* Variation of the *N*-acetyltransferase *2* gene in a Romanian and a Kyrgyz population. *Cancer Epidemiol Biomarkers Prev* 2006; **15:**138-141.

[188] Teixeira RLF, Miranda AB, Pacheco AG, Lopes MQP, Fonseca-Costa J, Rabahi MF *et al.* Genetic profile of the arylamine *N*-acetyltransferase 2 coding gene among individuals from two different regions of Brazil. *Mutat Res* 2007; **624:**31-40.

[189] Woolhouse NM, Qureshi MM, Bastaki SM, Patel M, Abdulrazzaq Y, Bayoumi RA. Polymorphic N-acetyltransferase (NAT2) genotyping of Emiratis. *Pharmacogenetics* 1997; **7:**73-82.

[190] Patin E, Barreiro LB, Sabeti PC, Austerlitz F, Luca F, Sajantila A *et al.* Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. *Am J Hum Genet* 2006; **78:**423-436.

[191] Dandara C, Masimirembwa CM, Magimba A, Kaaya S, Sayi J, Sommers dK *et al.* Arylamine N-acetyltransferase (NAT2) genotypes in Africans: the identification of a new allele with nucleotide changes 481C>T and 590G>A. *Pharmacogenetics* 2003; **13:**55-58.

[192] Delomenie C, Sica L, Grant DM, Krishnamoorthy R, Dupret JM. Genotyping of the polymorphic N-acetyltransferase (NAT2*) gene locus in two native African populations. *Pharmacogenetics* 1996; **6:**177-185.

[193] Bayoumi RA, Qureshi MM, al Ameri MM, Woolhouse NM. The N-acetyltransferase G191 A mutation among Sudanese and Somalis. *Pharmacogenetics* 1997; **7:**397-399.

[194] Al-Yahyaee S, Gaffar U, Al-Ameri MM, Qureshi M, Zadjali F, Ali BH *et al.* N-acetyltransferase polymorphism among northern Sudanese. *Hum Biol* 2007; **79:**445-452.

[195] Smith CAD, Wadelius M, Gough AC, Harrison DJ, Wolf CR, Rane A. A simplified assay for the arylamine N-acetyltransferase 2 polymorphism validated by phenotyping with isoniazid. *J Med Genet* 1997; **34:**758-760.

[196] Agúndez JA, Martínez C, Olivera M, Ledesma MC, Ladero JM, Benítez J. Molecular analysis of the arylamine N-acetyltransferase polymorphism in a Spanish population. *Clin Pharmacol Ther* 1994; **56:**202-209.

[197] Lemos MC, Regateiro FJ. N-acetyltransferase genotypes in the Portuguese population. *Pharmacogenetics* 1998; **8:**561-564.

[198] Lu JF, Cao XM, Liu ZH, Cao W, Guo LQ, Zhuo HT *et al.* Genetic analysis of N-acetyltransferase polymorphism in a Chinese population. *Zhongguo Yao Li Xue Bao* 1998; **19:**347-351.

[199] Tang TS, Jin SK, Lee JE, Woo SW, Roh J. Comparison of genetic polymorphisms of the *NAT2* gene between Korean and four other groups. *J Clin Pharm Ther* 2009; **34:**709-718.

[200] Singh N, Dubey S, Chinnaraj S, Golani A, Maitra A. Study of *NAT2* gene polymorphisms in an Indian population: Association with plasma isoniazid concentration in a cohort of tuberculosis patients. *Molecular Diagnosis & Therapy* 2009; **13:**49-58.

[201] Kukongviriyapan V, Prawan A, Tassaneyakul W, Aiemsa-Ard J, Warasiha B. Arylamine N-acetyltransferase-2 genotypes in the Thai population. *Br J Clin Pharmacol* 2003; **55:**278-281.

[202] Jorge-Nebert LF, Eichelbaum M, Griese EU, Inaba T, Arias TD. Analysis of six SNPs of NAT2 in Ngawbe and Embera Amerindians of Panama and determination of the Embera acetylation phenotype using caffeine. *Pharmacogenetics* 2002; **12:**39-48.

[203] Bailliet G, Santos MR, Alfaro EL, Dipierri JE, Demarchi DA, Carnese FR *et al.* Allele and genotype frequencies of metabolic genes in Native Americans from Argentina and Paraguay. *Mutation Research* 2007; **627:**171-177.

[204] Bakayev VV, Mohammadi F, Bahadori M, Sheikholslami M, Javeri A, Masjedi MR *et al.* Arylamine N-acetyltransferase 2 slow acetylator polymorphisms in unrelated Iranian individuals. *Eur J Clin Pharmacol* 2004; **60:**467-471.

[205] Adams CH, Werely CJ, Victor TC, Hoal EG, Rossouw G, van Helden PD. Allele frequencies for glutathione S-transferase and N-acetyltransferase 2 differ in African population groups and may be associated with oesophageal cancer or tuberculosis incidence. *Clin Chem Lab Med* 2003; **41:**600-605.

[206] Mark LC, Kayden HJ, Steele JM, Cooper JR, Berlin I, Rovenstine EA *et al.* The physiological disposition and cardiac effects of procaine amide. *The Journal of Pharmacology and Experimental Therapeutics* 1951; **102:**5-15.

[207] Fretland AJ, Doll MA, Gray K, Feng Y, Hein DW. Cloning, sequencing, and recombinant expression of NAT1, NAT2, and NAT3 derived from the C3H/HeJ (rapid) and A/HeJ (slow) acetylator inbred mouse: functional characterization of the activation and deactivation of aromatic amine carcinogens. *Toxicol Appl Pharmacol* 1997; **142:**360-366.

[208] Smelt VA, Mardon HJ, Sim E. Placental expression of arylamine N-acetyltransferases: evidence for linkage disequilibrium between NAT1*10 and NAT2*4 alleles of the two human arylamine N-acetyltransferase loci NAT1 and NAT2. *Pharmacol Toxicol* 1998; **83:**149-157.

[209] Henning S, Cascorbi I, Munchow B, Jahnke V, Roots I. Association of arylamine N-acetyltransferases NAT1 and NAT2 genotypes to laryngeal cancer risk. *Pharmacogenetics* 1999; **9:**103-111.

[210] World Health Organization,GS. World Health Organization: Treatment of Tuberculosis: Guidelines for national programmes. 2003. pp. 27-38.

[211] Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. *Am J Respir Crit Care Med* 2003; **167:**1472-1477.

[212] Huang YS, Chern HD, Su WJ, Wu JC, Chang SC, Chiang CH *et al.* Cytochrome P450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. *Hepatology* 2003; **37:**924-930.

[213] Kinzig-Schippers M, Tomalik-Scharte D, Jetter A, Scheidel B, Jakob V, Rodamer M *et al.* Should we use *N*-acetyltransferase type 2 genotyping to personalize isoniazid doses? *Antimicrob Agents Chemother* 2005; **49:**1733-1738.

[214] Peloquin CA, Jaresko GS, Young C-L, Keung AC, Bulpitt AE, Jeliffe RW. Population pharmacokinetic modeling of isoniazid, rifampicin amd pyrazinamide. *Antimicrob Agents Chemother* 1997; **41:**2670-2679.

[215] Peloquin CA. Therapeutic drug monitoring in the treatment of tuberculosis. *Drugs* 2002; **62:**2169-2183.

[216] Conte JE, Golden JA, McQuitty M, Kipps J, Duncan S, McKenna E *et al.* Effects of gender, AIDS, and acetylator status on intrapulmonary concentrations of isoniazid. *Antimicrob Agents Chemother* 2002; **46:**2358-2364.

[217] Weiner M, Burman W, Vernon A, Benator D, Peloquin CA, Khan A *et al.* Low isoniazid concentrations and outcome of tuberculosis treatment with once-weekly isoniazid and rifapentine treatment. *Am J Respir Crit Care Med* 2003; **167:**1341-1347.

[218] Burman W, Benator D, Vernon A, Khan A, Jones B, Silva C *et al.* Acquired rifamycin resistance with twice-weekly treatment of HIV-related tuberculosis. *Am J Respir Crit Care Med* 2006; **173:**350-356.

[219] Saukkonen JJ, Cohn DL, Jasmer RM, Schenker S, Jereb JA, Nolan CM *et al.* An Official ATS statement: Hepatotoxicity of Antituberculosis Therapy. *Am J Respir Crit Care Med* 2006; **174:**935-952.

[220] Sharma SK, Balamurugan A, Saha PK, Pandey RM, Mehra NK. Evaluation of clinical and immunogenetic risk factors for the development of hepatotoxicity during antituberculosis treatment. *Am J Respir Crit Care Med* 2002; **166:**916-919.

[221] Mitchell JR, Thorgeirsson UP, Black M, Timbrell JA, Snodgrass WR, Potter WZ *et al.* Increased incidence of isoniazid hepatitis in rapid acetylators:possible relation to hydrazine metabolites. *Clin Pharmacol Ther* 1975; **18:**70-79.

[222] Ohno M, Yamaguchi I, Yamamoto I, Fukuda T, Yokota S, Maekura R *et al.* Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. *Int J Tuberc Lung Dis* 2000; **4:**256-261.

[223] Mitchison DA. Assessment of new sterilising drugs for treating pulmonary tuberculosis by culture at 2 months. *Am Rev Resp Dis* 1993; **147:**1062-1063.

[224] Roy B. Anti-tuberculosis drug induced hepatotoxicity and genetic polymorphisms in drug-matabolising enzymes. *Current Pharmacogenomics* 2003; **1:**101-109.

[225] Tostmann A, Boeree MJ, Aarnoutse RE, de Lange WCM, van der Ven AJAM, Dekhuijzen R. Antituberculosis drug-induced hepatotoxicity: Concise up-to-date review. *J Gastroenterol Hepatol* 2008; **23:**192-202.

[226] Girling DJ. The hepatic toxicity of antituberculosis regimens containing isoniazid rifampicin and pyrazinamide. *Tubercle* 1978; **59:**13-32.

[227] Ellard GA, Mitchison DA, Girling DJ, Nunn AJ, Fox W. The hepatoxicity of Isoniazid among rapid and slow acetylators of the drug. *Am Rev Resp Dis* 1978; **118:**628-629.

[228] Sharma SK. Antituberculosis drugs and hepatotoxicity. *Infection, Genetics and Evolution* 2004; **4:**167-170.

[229] Yew WW, Leung CC. Antituberculosis drugs and hepatotoxicity. *Respirology* 2006; **11:**699-707.

[230] Mitchell JR, Zimmerman HJ, Ishak KG, Thorgeirsson UP, Timbrell JA, Snodgrass WR *et al.* Isoniazid liver injury: clinical spectrum, pathology, and probable pathogenesis. *Ann Intern Med* 1976; **84:**181-192.

[231] Lauterburg BH, Smith CV, Todd EL, Mitchell JR. Pharmacokinetics of the toxic hydrazino metabolites formed from isoniazid in humans. *J Pharmacol Exp Ther* 1985; **235:**566-570.

[232] Askgaard DS, Wilcke T, Dossing M. Hepatotoxicity caused by the combined action of isoniazid and rifampicin. *Thorax* 1995; **50:**213-214.

[233] Shen C, Meng Q, Zhang G, Hu W. Rifampicin exacerbates isoniazid-induced toxicity in human but not rat hepatocytes in tissue-like cultures. *Br J Pharmacol* 2008; **153:**784-791.

[234] Roy PD, Majumder M, Roy B. Pharmacogenomics of anti-TB drugs-related hepatotoxicity. *Pharmacogenomics* 2008; **9:**311-321.

[235] Facino RM, Carini M. The inhibitory effect of pyrazinamide on microsomal monooxygenase activities is related to the binding to reduced cytochrome P-450. *Pharmacol Res Commun* 1980; **12:**523-537.

[236] Nishimura Y, Kurata N, Sakurai E, Yasuhara H. Inhibitory effect of antituberculosis drugs on human cytochrome P450-mediated activities. *J Pharmacol Sci* 2004; **96:**293-300.

[237] Tostmann A, Boeree MJ, Peters WHM, Roelofs HMJ, Aarnoutse RE, van der Ven AJAM *et al.* Isoniazid and its toxic metabolite hydrazine induce in vitro pyrazinamide toxicity. *Int J Antimicrob Agents* 2008; **31:**577-580.

[238] Schaberg T, Rebhan K, Lode H. Risk factors for side-effect of isoniazid, rifampicin and pyrazinamide in patients hospitalised for pulmonary tuberculosis. *Eur Respir J* 1996; **9:**2026-2030.

[239] Pande JN, Pande A, Singh SP. Acetylator status, drug metabolism and disease. *Natl Med J India* 2003; **16:**24-26.

[240] Shakya R, Rao BS, Shrestha B. Incidence of hepatotoxicity due to antituberculosis medicines and assessment of risk factors. *Ann Pharmacother* 2004; **38:**1074-1079.

[241] Pande JN, Singh SPN, Khilnani GC, Khilnani S, Tandon RK. Risk factors for hepatotoxicity from antituberculosis drugs: a case-control study. *Thorax* 1996; **51:**132-136.

[242] Ellard GA, Gammon PT. Pharmacokinetics of isoniazid metabolism in man. *J Pharmacokinet Biopharm* 1976; **4:**83-113.

[243] Higuchi N, Tahara N, Yanagihara K, Fukushima K, Suyama N, Inoue Y *et al.* *NAT2*6A*, a haplotype of the *N*-acetyltransferase 2 gene, is an important biomarker for risk of anti-tuberculosis drug-induced hepatotoxicity in Japanese patients with tuberculosis. *World J Gastroenterol* 2007; **13:**6003-6008.

[244] Shell,RCH. Children of Bondage: A social history of the slave society at the Cape of Good Hope, 1652 - 1838. Johannesburg.: Witwatersrand University Press.; 1994.

[245] de Wit E, Delport W, Rugamika CE, Meintjes A, Möller M, van Helden PD *et al.* Genome-wide analysis of the structure of the South African Coloured population in the Western Cape. *Hum Genet* 2010; **128:**145-153.

[246] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* 1988; **16:**1215.

[247] Sambrook,J, Fritsch,EF, Maniatis,T. Molecular Cloning: A laboratory manual.Second edition. U.S.A.: Cold Spring Harbor Laboratory Press; 1989.

[248] Podhajska AJ, Szybalski W. Conversion of the *Fok*I endonuclease to a universal restriction enzyme: cleavage of phage M13mp7 DNA at predetermined sites. *Gene* 1985; **40:**169-182.

[249] Ausubel,FM, Brent,R, Kingston,RE, Moore,DD, Seidman,JG, Smith,JA *et al.* Current Protocols in Molecular Biology.John Wiley & Sons, New York.; 1987.

[250] Bassam BJ, Caetano-Anollés G, Gresshoff PM. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* 1991; **196:**80-83.

[251] Hamajima N, Saito T, Matsua K, Takahashi T, Tajima K. Polymerase chain reaction with confronting two-pair primers for polymorphism genotyping. *Japanese Journal of Cancer research* 2000; **91:**865-868.

[252] Innis,MA, Gelfand,DH. Optimization of PCRs. In PCR Protocols: A guide to methods and applications.1990 edition. London: Academic Press; 1990.

[253] Weber WW. Populations and genetic polymorphisms. *Mol Diagn* 1999; **4:**299-307.

[254] Tost JR, Vidal R, Caylá J, Díaz-Cabanela D, Jiménez A, Broquetas JM *et al.* Severe hepatotoxicity due to anti-tuberculosis drugs in Spain. *Int J Tuberc Lung Dis* 2005; **9:**534-540.

[255] Khalili H, Dashti-Khavidaki S, Amini M, Mahjub R, Hajiabdolbaghi M. Is there any difference between acetylator phenotypes in tuberculosis patients and healthy subjects? *Eur J Clin Pharmacol* 2010; **66:**261-267.

[256] Zaid RB, Nargis M, Neelotpol S, Sayeed MA, Banu A, Shurovi S *et al.* Importance of acetylator phenotype in the identity of Asian populations. *Hum Biol* 2007; **79:**363-368.

[257] Marzuki OA, Fauzi ARM, Kamarul Imran M. Prevalence and risk factors of anti-tuberculosis drug-induced hepatitis in Malaysia. *Singapore Med J* 2008; **49:**688-693.

[258] Lee EJ. Genetic polymorphisms in drug metabolism - its relevance to Asian populations. *Ann Acad Med Singapore* 1991; **20:**56-60.

[259] den Boon S, van Lill SWP, Borgdorff MW, Enarson DA, Verver S, Bateman ED *et al.* High prevalence of tuberculosis in previously treated patients, Cape Town, South Africa. *Emerging Infectious Diseases* 2007; **13:**1189-1194.

[260] Verver S, Warren RM, Beyers N, Richardson M, van der Spuy GD, Borgdorff MW *et al.* Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care Med* 2005; **171:**1430-1435.

[261] den Boon S, Verver S, Marais BJ, Enarson DA, Lombard CJ, Bateman ED *et al.* Association between passive smoking and infection with Mycobacterium tuberculosis in children. *Pediatrics* 2007; **119:**734-739.

[262] Kritzinger FE, den Boon S, Verver S, Enarson DA, Lombard CJ, Borgdorff MW *et al.* No decrease in annual risk of tuberculosis infection in endemic area in Cape Town, South Africa. *Trop Med Int Health* 2009; **14:**136-142.

[263] Quintana-Murci L, Harmant C, Quach H, Balanovsky O, Zaporozhchenko V, Bormans C *et al.* Strong maternal Khoisan contribution to the South African Coloured population: a case of gender-biased admixture. *American Journal of Human Genetics* 2010; **86:**611-620.

[264] Hoffmeyer S, Burk O, von Richter O, Arnold HP, Brockmöller J, Johne A *et al.* Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlations of one allele with P-glycoprotein expression and activity in vivo. *Proc Natl Acad Sci U S A* 2000; **97:**3473-3478.

[265] Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV *et al.* A "silent" polymorphism in the *MDR*1 gene changes substrate specificity. *Science* 2007; **315:**525-528.

[266] Agúndez JAG, Golka K, Martínez C, Selinski S, Blaszkewicz M, Garcia-Martin E. Unraveling ambiguous NAT2 genotyping data. *Clin Chem* 2008; **54:**1390-1394.

[267] Abe M, Suzuki T, Deguchi T. An improved method for genotyping of N-acetyltransferase polymorphism by polymerase chain reaction. *Jpn J Hum Genet* 1993; **38:**163-168.

[268] O'Neil WM, Drobitch RK, MacArthur RD, Farrough MJ, Doll MA, Fretland AJ *et al.* Acetylator phenotype and genotype in patients infected with HIV: discordance between methods for phenotype determination and genotype. *Pharmacogenetics* 2000; **10:**171-182.

[269] Bolt HM, Selinski S, Dannappel D, Blaszkewicz M, Golka K. Re-investigation of the concordance of human NAT2 phenotypes and genotypes. *Arch Toxicol* 2005; **79:**196-200.

[270] Golka K, Blaszkewicz M, Samimi M, Bolt HM, Selinski S. Reconstruction of N-acetyltransferase 2 haplotypes using PHASE. *Arch Toxicol* 2008; **82:**265-270.

[271] Agúndez JAG. NAT2 Genotyping: Equilibrium between accuracy and feasibility in routine analysis. *The Journal of Applied Research* 2003; **3:**1-6.

[272] Deitz AC, Rothman N, Rebbeck TR, Hayes RB, Chow WH, Zheng W *et al.* Impact of misclassification in genotype-exposure interaction studies: example of N-acetyltransferase 2 (NAT2), smoking, and bladder cancer. *Cancer Epidemiol Biomarkers Prev* 2004; **13:**1543-1546.

[273] Sabbagh A, Darlu P. Inferring haplotypes at the *NAT2* locus: the computational approach . *BMC Genetics* 2005; **6:**30.

[274] Sabbagh A, Darlu P. SNP selection at the *NAT2* locus for an accurate prediction of the acetylation phenotype. *Genetics in Medicine* 2006; **8:**76-85.

[275] Luca F, Bubba G, Basile M, Brdicka R, Michalodimitrakis E, Rickards O *et al.* Multiple advantageous amino acid variants in the NAT2 gene in human populations. *PLoS ONE* 2008; **3:**e3136.

[276] Thorp,CR. Hunter-gatherers and farmers: an enduring frontier in the Caledon Valley, South Africa. Oxford, England: Archaeopress, Publishers of British Archaeological Reports; 2000.

[277] Magalon H, Patin E, Austerlitz F, Hegay T, Aldashev A, Quintana-Murci L *et al.* Population genetic diversity of the *NAT2* gene supports a role of acetylation in human adaptation to farming in Central Asia. *Eur J Hum Genet* 2008; **16:**243-251.

[278] Sabbagh A, Darlu P, Crouau-Roy B, Poloni ES. Arylamine *N*-acetyltransferase 2 (*NAT2*) genetic diversity and traditional subsistence: A worldwide population survey. *PLoS ONE* 2011; **6:**e18507.

[279] Grant DM, Tang BK, Kalow W. Polymorphic N-acetylation of a caffeine metabolite. *Clin Pharmacol Ther* 1983; **33:**355-359.

[280] Vincent-Viry M, Pontes ZB, Gueguen R, Galteau MM, Siest G. Segregation analyses of four urinary caffeine metabolite ratios implicated in the determination of human acetylation phenotypes. *Genet Epidemiol* 1994; **11:**115-129.

[281] Miners JO, Birkett DJ. The use of caffeine as a metabolic probe for human drug metabolizing enzymes. *Gen Pharmacol* 1996; **27:**245-249.

[282] Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics* 2008; **40:**575-583.

[283] Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science* 2009; **322:**881-888.

[284] Hardy J, Singleton A. Genomewide association studies in human disease. *N Engl J Med* 2009; **360:**1759-1768.

[285] Verver S, Warren R, Munch Z, Vynnycky E, van Helden PD, Richardson M *et al.* Transmission of tuberculosis in a high incidence urban community in South Africa. *Int J Epidemiol* 2004; **33:**351-357.

[286] Shisana, O., Rehle, T., Simbayi, L. C., Zuma, K., Jooste, S., Pillay-van Wyk, V. *et al*. South African National HIV prevalence, incidence, behaviour and communication survey, 2008. http://www.hsrc.ac.za/Document-3238.phtml#. Downloaded 13 June 2011 12:29:04. 2009.
Ref Type: Report

[287] Nurse,GT, Weiner,JS, Jenkins,T. The people of Southern Africa and their affinities. Oxford, U.K.: Clarendon Press; 1985.

[288] Erasmus,Z. Coloured by history, shaped by place: new perspectives on coloured identities in Cape Town.First edition.Kwela books; 2001.

[289] Adhikari,M. Not white enough, not black enough: racial identity in the South African coloured community.Ohio University Press; 2005.

[290] Barreiro LB, Neyrolles O, Babb CL, Tailleux L, Quach H, McElreavey K *et al.* Promoter variation in the DC-SIGN-encoding gene CD*209* is associated with tuberculosis. *PLoS Medicine* 2006; **3:**e20.

[291] Lombard Z, Brune AE, Hoal EG, Babb C, van Helden PD, Epplen JT *et al.* HLA class II disease associations in Souther Africa. *Tissue Antigens* 2006; **67:**97-110.

[292] Walzl G, Beyers N, van Helden P. TB: a partnership for the benefit of research and community. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2005; **995:**S15-S19.

[293] Brahmbhatt S, Black GF, Carroll NM, Beyers N, Salker F, Kidd M *et al.* Immune markers measured before treatment predict outcome of intensive phase tuberculosis therapy. *Clinical and Experimental Immunology* 2006; **146:**243-252.

[294] Lienhardt C. From exposure to disease: the role of environmental factors in susceptibility to and development of tubecrulosis. *Epidemiological Reviews* 2001; **23:**288-301.

[295] Bellamy R. Susceptibility to mycobacterial infections: The importance of host genetics. *Genes Immun* 2003; **4:**4-11.

[296] Blackwell JM. Genetics and genomics in infectious disease susceptibility. *Trends Mol Med* 2001; **7:**521-526.

[297] Rezaei N, Aghamohammadi A, Mansouri D, Parvaneh N, Casanova J-L. Tuberculosis: A new look at an old disease. *Expert Rev Clin Immunol* 2011; **7:**129-131.

[298] Warren R, Hauman J, Beyers N, Richardson M, Schaaf HS, Donald P *et al.* Unexpectedly high strain diversity of *Mycobacterium tuberculosis* in a high-incidence community. *S Afr Med J* 1996; **86:**45-49.

[299] Warren R, Richardson M, van der Spuy G, Victor T, Sampson S, Beyers N *et al.* DNA fingerprinting and molecular epidemiology of tuberculosis: Use and interpretation in an epidemic setting. *Electrophoresis* 1999; **20:**1807-1812.

[300] Babb C, van der Merwe L, Beyers N, Pheiffer C, Walzl G, Duncan K *et al.* Vitamin D receptor gene polymorphisms and sputum conversion time in pulmonary tuberculosis patients. *Tuberculosis* 2007; **87:**295-302.

[301] Enarson,DA, Rieder,HL, Arnadottir,T, Trébucq,A. Management of tuberculosis: a guide for low income countries.Fifth edition. Paris: International Union Against Tuberculosis and Lung Disease; 2000.

[302] Hesseling AC, Walzl G, Enarson DA, Carroll NM, Duncan K, Lukey PT *et al.* Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *Int J Tuberc Lung Dis* 2010; **14:**560-570.

[303] Walraven JM, Zang Y, Trent JO, Hein DW. Structure/Function evaluations of single nucleotide polymorphisms in human N-acetyltransferase 2. *Curr Drug Metab* 2008; **9:**471-486.

[304] Grantham R, Gautier C, Gouy M, Jacobzone M, Mercier R. Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Research* 1981; **9:**r43-r74.

[305] Eyre-Walker AC. An analysis of codon usage in mammals: selection or mutation bias? *Journal of Molecular Evolution* 1991; **33:**442-449.

[306] Plotkin JB, Robins H, Levine AJ. Tissue-specific codon usage and the expression of human genes. *PNAS* 2004; **101:**12588-12591.

[307] Bennetzen JL, Hall BD. Codon selection in yeast. *J Biol Chem* 1982; **257:**3026-3031.

[308] Gouy M, Gautier C. Codon usage in bacteria: Correlation with gene expressivity. *Nucleic Acids Research* 1982; **10:**7055-7074.

[309] Sharp PM, Tuohy TM, Mosurski KR. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Research* 1986; **14:**5125-5143.

[310] Wang D, Johnson AD, Papp AC, Kroetz DL, Sadée W. Multidrug resistance polypeptide 1 (*MDR1*, *ABCB1*) variant 3435C>T affects mRNA stability. *Pharmacogenet Genomics* 2005; **15:**693-704.

[311] Werely CJ, Donald PR, van Helden PD. *NAT2* polymorphisms and their influence on the pharmacology and toxicology of isoniazid in TB patients. *Personalised Medicine* 2007; **4:**123-131.

[312] Garcia-Closas M, Hein DW, Silverman D, Malats N, Yeager M, Jacobs K *et al.* A single nucleotide polymorphism tags variation in the arylamine *N*-acetyltransferase 2 phenotype in populations of European background. *Pharmacogenet Genomics* 2011; **21:**231-236.

[313] Martínez C, Agúndez JA, Olivera M, Martín R, Ladero JM, Benítez J. Lung cancer and mutations at the polymorphic NAT2 gene locus. *Pharmacogenetics* 1995; **5:**207-214.

[314] Agúndez JA, Olivera M, Ladero JM, Rodriguez-Lescure A, Ledesma MC, Diaz-Rubio M *et al.* Increased risk for hepatocellular carcinoma in NAT2-slow acetylators and CYP2D6-rapid metabolizers. *Pharmacogenetics* 1996; **6:**501-512.

[315] Sabbagh A, Langaney A, Darlu P, Gerard N, Krishnamoorthy R, Poloni E. Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. *BMC Genetics* 2008; **9:**21.

[316] Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21:**263-265.

[317] Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am J Hum Genet* 2005; **76:**449-462.

[318] Richardson M, van Lill SWP, van der Spuy GD, Munch Z, Booysen CN, Beyers N *et al.* Historic and recent events contribute to the disease dynamics of Beijing-like *Mycobacterium tuberculosis* isolates in a high incidence region. *Int J Tuberc Lung Dis* 2002; **6:**1001-1011.

[319] Marais BJ, Victor TC, Hesseling AC, Barnard M, Jordaan A, Brittle W *et al.* Beijing and Haarlem genotypes are overrepresented among children with drug-resistant tuberculosis in the Western Cape province of South Africa. *J Clin Microbiol* 2006; **44:**3539-3543.

[320] Parwati I, van Crevel R, Sudiro M, Alisjahbana B, Pakasi T, Kremer K *et al.* *Mycobacterium tuberculosis* population structures differ significantly on two Indonesian islands. *J Clin Microbiol* 2008; **46:**3639-3645.

[321] Parwati I, Alisjahbana B, Apriani L, Soetikno RD, Ottenhoff TH, van der Zanden AGM *et al. Mycobacterium tuberculosis* Beijing genotype is an independant risk factor for tuberculosis treatment failure in Indonesia. *J Infect Dis* 2010; **201:**553-557.

[322] Hoal EG, Lewis LA, Jamieson SE, Tanzer F, Rossouw M, Victor T *et al.* SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. *Int J Tuberc Lung Dis* 2004; **8:**1464-1471.

[323] van Crevel R, Parwati I, Sahiratmadja E, Marzuki S, Ottenhoff THM, Netea MG *et al.* Infection with *Mycobacterium tuberculosis* Beijing genotype strains is associated with polymorphisms in *SLC11A1/NRAMP1* in Indonesian patients with tuberculosis. *J Infect Dis* 2009; **200:**1671-1674.

[324] Ulrichs T, Kaufmann SH. New insights into the function of granulomas in human tuberculosis. *J Pathol* 2006; **208:**261-269.

[325] Russell DG, Cardona P-J, Kim M-J, Allain S, Altare F. Foamy macrophages and the progression of the human tuberculosis granuloma. *Nature Immunology* 2009; **10:**943-948.

[326] Russell DG, Barry CE3, Flynn JL. Tuberculosis: What we don't know can, and does , hurt us. *Science* 2010; **328:**852-856.

[327] Tsai CJ, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R. Synonymous mutations and robosome stalling can lead to altered folding pathways and distinct minima. *J Mol Biol* 2008; **383:**281-291.

[328] Lee WM. Drug-induced hepatotoxicity. *N Engl J Med* 1995; **333:**1118-1127.

[329] Lee SW, Chung LS-C, Huang HH, Chuang T-Y, Liou Y-H, Wu LS-H. *NAT2* and *CYP2E1* polymorphisms and susceptibility to first-line anti-tuberculosis drug-induced hepatitis. *Int J Tuberc Lung Dis* 2010; **14:**622-6.

[330] Carroll NM, Uys P, Hesseling A, Lawrence K, Pheiffer C, Salker F *et al.* Prediction of delayed treatment response in pulmonary tuberculosis: Use of time to positivity values of Bactec cultures. *Tuberculosis* 2008; **88:**624-630.

[331] Gunawan B, Kaplowitz N. Clinical perspectives on xenobiotic-induced hepatotoxicity. *Drug Metab Rev* 2004; **36:**301-312.

[332] Lucena MI, Andrade RJ, Martínez C, Ulzurrun E, García-Martín E, Borraz Y *et al.* Glutathione S-Transferase M1 and T1 null genotypes increase susceptibility to idiosyncratic drug-induced liver injury. *Hepatology* 2008; **48:**588-596.

[333] Hein DW, McQueen CA, Grant DM, Goodfellow GH, Kadlubar FF, Weber WW. Pharmacogenetics of the arylamine N-acetyltransferases: a symposium in honor of Wendell W. Weber. *Drug Metab Dispos* 2000; **28:**1425-1432.

[334] Dutcher JS, Strong JM, Lucas SV, Lee WK, Atkinson AJJ. Procainamide and N-acetylprocainamide kinetics investigated simultaneously with stable isotope methodology. *Clin Pharmacol Ther* 1977; **22:**447-457.

[335] Cribb AE, Isbrucker R, Levatte T, Tsui B, Gillespie CT, Renton KW. Acetylator phenotyping: the urinary caffeine metabolite ratio in slow acetylators correlates with a marker of systemic NAT1 activity. *Pharmacogenetics* 1994; **4:**166-170.

[336] Tishkoff SA, Williams SM. Genetic analysis of African populations: Human evolution and complex disease. *Nature Reviews Genetics* 2002; **3:**611-621.

[337] Beckmann JS, Estivill X, Antonarakis SE. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nature Reviews Genetics* 2007; **8:**639-646.

[338] de Smith AJ, Walters RG, Froguel P, Blakemore AI. Human genes involved in copy number variation: mechanisms of origin, functional effects and implications for disease. *Cytogenet Genome Res* 2008; **123:**17-26.

[339] Badawi AF, Hirvonen A, Bell DA, Lang NP, Kadlubar FF. Role of aromatic amine acetyltransferases, NAT1 and NAT2, in carcinogen-DNA adduct formation in the human urinary bladder. *Cancer Res* 1995; **55:**5230-5237.

[340] de Leon JH, Vatsis KP, Weber WW. Characterization of naturally occurring and recombinant human N-acetyltransferase variants encoded by NAT1*. *Mol Pharmacol* 2000; **58:**288-299.

[341] Zhu Y, Hein DW. Functional effects of single nucleotide polymorphisms in the coding region of human *N*-acetyltransferase 1. *The Pharmacogenomics Journal* 2008; **8:**339-348.

[342] Fretland AJ, Doll MA, Leff MA, Hein DW. Functional characterization of nucleotide polymorphisms in the coding region of N-acetyltransferase 1. *Pharmacogenetics* 2001; **11:**511-520.

[343] Vaziri SA, Hughes NC, Sampson H, Darlington G, Jewett MA, Grant DM. Variation in enzymes of arylamine procarcinogen biotransformation among bladder cancer patients and control subjects. *Pharmacogenetics* 2001; **11:**7-20.

[344] Hein DW, Rustan TD, Ferguson RJ, Doll MA, Gray K. Metabolic activation of aromatic and heterocyclic N-hydroxyarylamines by wild-type and mutant recombinant human NAT1 and NAT2 acetyltransferases. *Arch Toxicol* 1994; **68:**129-133.

[345] Mortensen H, Froment A, Lema G, Bodo J-M, Ibrahim M, Nyambo TB *et al.* Characterization of genetic variation and natural selection at the arylamine *N*-acetyltransferase genes in global human populations. *Pharmacogenomics* 2011; **12:**1545-1558.

[346] Varzim G, Monteiro E, Silva R, Pinheiro C, Lopes C. Polymorphisms of arylamine N-acetyltransferase (NAT1 and NAT2) and larynx cancer susceptibility. *ORL J Otorhinolaryngol Relat Spec* 2002; **64:**206-212.

[347] Li D, Jiao L, Li Y, Doll MA, Hein DW, Bondy ML *et al.* Polymorphisms of cytochrome P4501A2 and N-acetyltransferase genes, smoking, and risk of pancreatic cancer. *Carcinogenesis* 2006; **27:**103-111.

[348] Lilla C, Verla-Tebit E, Risch A, Jager B, Hoffmeister M, Brenner H *et al.* Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. *Cancer Epidemiol Biomarkers Prev* 2006; **15:**99-107.

[349] Blum M, Demierre A, Grant DM, Heim M, Meyer UA. Molecular mechanism of slow acetylation of drugs and carcinogens in humans. *Proc Natl Acad Sci U S A* 1991; **88:**5237-5241.

[350] Vatsis KP, Martell KJ, Weber WW. Diverse point mutations in the human gene for polymorphic N- acetyltransferase. *Proc Natl Acad Sci U S A* 1991; **88:**6333-6337.

[351] Martinez C, Agundez JA, Olivera M, Martin R, Ladero JM, Benitez J. Lung cancer and mutations at the polymorphic NAT2 gene locus. *Pharmacogenetics* 1995; **5:**207-214.

[352] Lin HJ, Probst-Hensch NM, Hughes NC, Sakamoto GT, Louie AD, Kau IH *et al.* Variants of N-acetyltransferase NAT1 and a case-control study of colorectal adenomas. *Pharmacogenetics* 1998; **8:**269-281.

[353] Hein DW. N-acetyltransferase 2 genetic polymorphism: effects of carcinogen and haplotype on urinary bladder cancer risk. *Oncogene* 2006; **25:**1649-1658.

[354] Huang YS, Chern HD, Su WJ, Wu JC, Lai SL, Yang SY *et al.* Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor for antituberculosis drug-induced hepatitis. *Hepatology* 2002; **35:**883-889.

[355] Doll MA, Fretland AJ, Deitz AC, Hein DW. Determination of human NAT2 acetylator genotype by restriction fragment-length polymorphism and allele-specific amplification. *Anal Biochem* 1995; **231:**413-420.

[356] Blomeke B, Sieben S, Spotter D, Landt O, Merk HF. Identification of N-acetyltransferase 2 genotypes by continuous monitoring of fluorogenic hybridization probes. *Anal Biochem* 1999; **275:**93-97.

[357] Osborne A, Bell C, Grant F, Dick F, Seaton A, Haites N. A rapid method of screening for N-acetyltransferase (NAT2) phenotype by use of the WAVE DNA fragment analysis system. *Biochem Genet* 2003; **41:**405-411.

[358] Doll MA, Hein DW. Comprehensive Human NAT2 Genotype Method Using Single Nucleotide Polymorphism-Specific Polymerase Chain Reaction Primers and Fluorogenic Probes. *Anal Biochem* 2001; **288:**106-108.

[359] Bunschoten A, Tiemersma E, Schouls L, Kampman E. Simultaneous determination of polymorphism in N-acetyltransferase 1 and 2 genes by reverse line blot hybridization. *Anal Biochem* 2000; **285:**156-162.

[360] Simsek M, Tanira MO, Al Baloushi KA, Al Barwani HS, Lawatia KM, Bayoumi RA. Improved detection of the T(341)C polymorphism in human N-acetyltransferase (NAT2) gene by using two complementary PCR-RFLP methods: better than automated DNA sequencing for detection of heterozygotes. *Clin Biochem* 2000; **33:**585-588.

[361] Schaaf HS, Parkin DP, Seifart HI, Werely CJ, Hesseling PB, van Helden PD *et al.* Isoniazid pharmacokinetics in children treated for respiratory tuberculosis. *Archives of Disease in Childhood* 2005; **90:**614-618.

[362] Vallone PM, Butler JM. AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 2004; **37:**226-231.

[363] Bujalkova M, Zavodna K, Krivulcik T, Ilencikova D, Wolf B, Kovac M *et al.* Multiplex SNaPshot genotyping for detecting loss of heterozygosity in the mismatch-repair genes MLH1 and MSH2 in microsatellite-unstable tumors. *Clin Chem* 2008; **54:**1844-1854.

[364] Bray NJ, Buckland PR, Hall H, Owen MJ, O'Donovan MC. The serotonin-2A receptor gene locus does not contain common polymorphism affecting mRNA levels in adult brain. *Mol Psychiatry* 2004; **9:**109-114.

[365] Vallone PM, Just RS, Coble MD, Butler JM, Parsons TJ. A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. *Int J Legal Med* 2004; **118:**147-157.

[366] Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. *Nature Reviews Genetics* 2004; **5:**335-344.

[367] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005; **437:**376-380.

[368] Shendure J, Hanlee J. Next-generation DNA sequencing. *Nat Biotechnol* 2008; **26:**1135-1145.

[369] Healy K. Nanopore-based single-molecule DNA analysis. *Nanomedicine* 2007; **2:**459-481.

[370] Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA *et al.* Origins and affinities of modern humans: A comparison of mitochondrial and nuclear genetic data. *Am J Hum Genet* 1995; **57:**523-538.

[371] Melton T, Ginther C, Sensabaugh G, Soodyall H, Stoneking M. Extent of heterogeneity in mitochondrial DNA of sub-Saharan African populations. *J Forensic Sci* 1997; **42:**582-592.

[372] Vesell ES. Genetic and environmental factors affecting drug response in man. *Fed Proc* 1972; **31:**1253-1269.

[373] Mahgoup A, Idle JR, Dring LG, Lancaster R, Smith RL. Polymorphic hydroyxlation of debrisoquine in man. *Lancet* 1977; **310:**584-586.

[374] Zhong S, Wyllie AH, Barnes D, Wolf CR, Spurr NK. Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer. *Carcinogenesis* 1993; **14:**1821-1824.

[375] Landi S. Mammalian class theta GST and differential susceptibility to carcinogens: A review. *Mutat Res* 2000; **463:**247-283.

[376] Baquet CR, Commiskey P, Mack K, Meltzer S, Mishra S. Oesophageal cancer epiemiology in Blacks and Whites: racial and gender disparities in incidence, mortality, survival rates and histology. *Journal of the National Medical Association* 2005; **97:**1471-1478.

[377] Yang CS. Oesophageal cancer in China. *Cancer Res* 1980; **40:**2633-2644.

[378] Warwick GP. Some aspects of the epidemiology and aetiology of oesophageal cancer with particular emphasis on the Transkei, South Africa. *Advances in Cancer Research* 1973; **17:**81-229.

[379] Meyer,D, Liebenberg,P. Cataract Etiology: A comprehensive Review. In: Agarwal,S, Agarwal,A, Apple,DJ, Buratto,L, Alió,JL, Pandey,SK, Agarwal,A, editors. Textbook of Ophthalmology. New Delhi: Jaypee Brothers; 2002. pp. 1587-1619.

[380] Meyer D, Parkin DP, Seifart HI, Maritz JS, Engelbrecht AH, Werely CJ *et al.* NAT2 slow acetylator function as a risk indicator for age-related cataract formation. *Pharmacogenetics* 2003; **13:**285-289.

[381] Tamer L, Yilmaz A, Yildirim H, Ayaz L, Ates NA, Karakas S *et al. N*-acetyltransferase 2 phenotype may be associated with susceptibility to age related cataract. *Curr Eye Res* 2005; **30:**835-839.

[382] West S, Muñoz B, Schein OD, Vitale S, Maguire M, Taylor HR *et al.* Cigarette smoking and risk for progression of nuclear opacities. *Arch Ophthalmol* 1995; **113:**1377-1380.

[383] Dairou J, Malecaze F, Dupret JM, Rodrigues-Lima F. The xenobiotic-metabolizing enzymes arylamine N-acetyltransferases in human lens epithelial cells: inactivation by cellular oxidants and UVB-induced oxidative stress. *Mol Pharmacol* 2005; **67:**1299-1306.

[384] Weber WW, Hein DW. Clinical pharmacokinetics of isoniazid. *Clin Pharmacokinet* 1979; **4:**401-422.

[385] Hiratsuka M, Kishikawa K, Takekuma Y, Matsuura M, Narahara K, Inoue T *et al.* Genotyping of the *N*-acetylatransferase2 polymorphism in the prediction of adverse drug reactions to isoniazid in Japanese patients. *Drug Metabol Pharmacokin* 2002; **17:**357-362.

[386] Durbin RM, Abesasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA *et al.* A map of human genome variation from population-scale sequencing. *Nature* 2010; **467:**1061-1073.

[387] Vatsis KP, Weber WW. Structural heterogeneity of Caucasian N-acetyltransferase at the NAT1 gene locus. *Arch Biochem Biophys* 1993; **301:**71-76.

[388] Sim E, Westwood I, Fullam E. Arylamine N-acetyltransferases. *Expert Opin Drug Metab Toxicol* 2007; **3:**169-184.

Addendum

A1 Internet Resources and additional information

I. World Health Organization Report 2006: Global Tuberculosis Control – surveillance, planning, financing. Downloaded February 15, 2007. (http://www.who.int/tb/publications/global_report/en/index.html).

II. World Health Organization. Tuberculosis Fact Sheet No. 104 – Global and regional incidence (March, 2007). Downloaded 3 April, 2007. (http://www.who.int/tb/mediacentre/factsheet/fs104/en/index.html)

III. The International Programs Centre , part of the Population Division of the U.S. Census Bureau. Downloaded 28 August, 2008. (http://www.census.gov/ipc.www).

IV. World Health Organization Report 2008: Global Tuberculosis Control – surveillance, planning, financing. Downloaded August 2008. (http://www.who.int/tb/publication/global_report/2008/en/)

V. Health Systems Trust. Downloaded 26 May 2008. (http://www.hst.org.za/healthstats/161/data).

VI. World Lung Foundation. Tuberculosis: South Africa. Downloaded august 2008. (http://www.worldlungfoundation.org/map_SA.php)

VII. Human *NAT1* Gene Sequence. (http://www.ncbi.nlm.nih.gov/nuccore/51466871 and http://www.ncbi.nlm.nih.gov/nuccore/14018355).

VIII. Consensus Human aArylamine *N*-Acetyltransferase Gene Nomenclature. (http://www.louisville.edu/medschool/pharmacology/consensus-human-arylamine-n-acetyltransferase-gene-nomenclature/)

IX. D.P. Parkin. Ph.D. Dissertation, 1996. Stellenbosch University (http://hdl.handle.net/10019.1/1267).

X. Human NAT2 Gene Sequence. (http://www.ncbi.nlm.nih.gov/nuccore/219871).

XI. Primer Premier Primer Design Programme. (http://www.PremierBiosoft.com)

XII. BLAST Sequence Alignment Algorithm. (http://www.nlm.nih.gov/blast/blast.cgi).

XIII. Sequencher Sequence Analysis Programme (http://www.genecodes.com).

XIV. Stob TB. (http://www.stoptb.org/countries/tbdata.asp).

XV. Statistical Release P0301. Community Survey, 2007 (revised version Downloaded 15 October 2010. (http://www.statssa.gov.za/publications/P0301/P0301.pdf).

XVI. Haploview, version 4.2 (http://www.broadinstitute.org/mpg/haploview).

XVII. SAS Statistical Analysis Programme (http://sas.com).

XVIII. Statistica, Statistical Analysis Programme (http://www.statsoft.com).

XIX. PHASE Programme (http://www.stat.washington.edu/stephens/software.html).

XX. C.L. Babb, Ph.D. Dissertation, 2007. Stellenbosch University (.

XXI. World health Organization Factsheet No. 104 (2010). Downloaded December, 2010. (http://www.who.int/mediacentre/factsheet/fs104/en/).

XXII. Integrated DNA Technologies. (http://www.idtdna.com).

XXIII. Homo sapiens *NAT1*4* Allele Sequence for the Arylamine N-Acetyltransferase 1 gene (http://www.ncbi.nlm.nih.gov/nuccore/AJ307007).