

INSTITUTO OSWALDO CRUZ  
Doutorado em Biologia Celular e Molecular

**Influência da variabilidade genética humana dos genes  
*NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* na ocorrência de reações  
adversas hepáticas induzidas pela isoniazida em pacientes  
com tuberculose ativa**

**RAQUEL LIMA DE FIGUEIREDO TEIXEIRA**

Rio de Janeiro

2009

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Ficha catalográfica elaborada pela  
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

T266

Teixeira, Raquel Lima de Figueiredo

Influência da variabilidade genética humana dos genes NAT2, CYP2E1, GSTM1, GSTT1 na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes tuberculose ativa / Raquel Lima de Figueiredo Teixeira. – Rio de Janeiro, 2009.

xvi, 205 f. : il. ; 30 cm.

Tese (doutorado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Celular e Molecular, 2009.

Bibliografia: f. 175-190

1. Tuberculose. 2. Enzimas metabolizadoras de drogas. 3. Farmacogenética. 4. Hepatotoxicidade. 5. População brasileira. I. Título.

CDD 616 995



Ministério da Saúde  
Fundação Oswaldo Cruz  
Instituto Oswaldo Cruz  
Curso de Pós-Graduação em Biologia Celular e Molecular

## **RAQUEL LIMA DE FIGUEIREDO TEIXEIRA**

Influência da variabilidade genética humana dos genes *NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes com tuberculose ativa

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular, na área de Genética Molecular Humana.

Orientadores: Dr. Adalberto Rezende Santos  
Dr. Antonio Basílio de Miranda

Este trabalho foi realizado no Laboratório de Biologia Molecular aplicada a Micobactérias e no Laboratório de Genética Humana, ambos do Instituto Oswaldo Cruz - FIOCRUZ

RIO DE JANEIRO

2009



Ministério da Saúde

Fundação Oswaldo Cruz

Instituto Oswaldo Cruz

Curso de Pós-Graduação em Biologia Celular e Molecular

**Influência da variabilidade genética humana dos genes *NAT2*, *CYP2E1*,  
*GSTM1* e *GSTT1* na ocorrência de reações adversas hepáticas induzidas  
pela isoniazida em pacientes com tuberculose ativa**

RAQUEL LIMA DE FIGUEIREDO TEIXEIRA

Orientadores: Dr. Adalberto Rezende Santos

Dr. Antonio Basílio de Miranda

Data da aprovação: 29 de abril de 2009

BANCA EXAMINADORA:

TITULARES:

Dra. Márcia Mattos Gonçalves Pimentel - UERJ

Dr. Marcus Barreto Conde- UFRJ

Dr. Marcelo Alves Ferreira - FIOCRUZ

SUPLENTE:

Dr. Harrison Magdinier Gomes - FIOCRUZ

Dr. Pedro Hernan Cabello - FIOCRUZ (revisor)

RIO DE JANEIRO, 29 DE ABRIL DE 2009

## INSTITUTO OSWALDO CRUZ

“Influência da variabilidade genética humana dos genes *NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes com tuberculose ativa”

### RESUMO

Tese de doutorado

Raquel Lima de F. Teixeira

A isoniazida (INH), uma das principais drogas usadas no esquema de tratamento de primeira linha anti-tuberculose (anti-TB), tem sido associada à intensificação de efeitos adversos principalmente hepáticos. Sabe-se hoje, que variações polimórficas em genes humanos codificando enzimas envolvidas na biotransformação de diferentes fármacos podem contribuir para diferenças interindividuais na resposta farmacológica ou toxicológica de várias drogas estando diretamente relacionadas aos desfechos de falha terapêutica e reações adversas a drogas (ADRs). A INH é geralmente administrada de forma oral sendo metabolizada por enzimas hepáticas. O mecanismo de lesão hepatocelular induzido pela isoniazida envolve as enzimas de biotransformação N-acetiltransferase 2 e a mono-oxigenase CYP450 2E1. Mutações pontuais na região codificante de *NAT2* são capazes de alterar a atividade de acetilação da enzima gerando três possíveis fenótipos: acetiladores lentos, intermediários e rápidos. Os indivíduos que apresentam uma acetilação lenta acumulam intermediários hepatotóxicos levando a ocorrência de ADRs. Outra classe de enzimas cuja participação na biotransformação da isoniazida vem sendo especulada é a glutathione S-transferase (GST). Considerando a importância das variantes alélicas dos genes envolvidos na metabolização da isoniazida e que suas frequências variam entre as diferentes etnias este trabalho teve como objetivos principais: (i) análise descritiva da distribuição dos alelos de *NAT2* em duas regiões diferentes do Brasil; (ii) identificação de novas mutações e caracterização haplotípica de novos alelos de *NAT2* bem como análise estrutural da influência dessas alterações na estrutura protéica de *NAT2* e (iii) estudo de associação, caso-controle, entre as variáveis genéticas presentes nos genes que codificam para *NAT2*, *CYP2E1* e GSTs humanas, com a ocorrência de reações adversas em pacientes com TB em tratamento com esquemas contendo isoniazida. Dezesete SNPs previamente descritos foram identificados na população estudada, dos quais, sete: 191G>A; 282C>T; 341T>C; 481C>T; 590G>A; 803A>G e 857G>A são os mais frequentes na população mundial. Adicionalmente, seis mutações novas foram identificadas e sete novos alelos de *NAT2* circulantes no Rio de Janeiro e/ou Goiás foram caracterizados através de clonagem e resequenciamento e experimentos de modelagem molecular. Nossos resultados mostraram a predominância de alelos de *NAT2* associados com o fenótipo de acetilação lenta em indivíduos brasileiros e que a distribuição desses alelos varia significativamente de acordo com a região brasileira estudada. Além disso, fomos capazes de constatar que os acetiladores lentos apresentaram um risco significativamente maior em desenvolver hepatotoxicidade (OR: 2,62; IC 95%: 1,75-3,49;  $p = 0,03$ ) ou hepatite medicamentosa (OR: 3,59; IC 95%: 2,53-4,64;  $p = 0,02$ ) quando comparados aos acetiladores intermediários/rápidos. Por outro lado, não observamos qualquer relação entre polimorfismos nos genes *CYP2E1*, *GSTM1* e *GSTT1* com a ocorrência de ADRs durante tratamento anti-TB. Sendo assim, nossos resultados sugerem a influência de variantes gênicas de *NAT2* na ocorrência de hepatotoxicidade durante o tratamento da tuberculose em indivíduos brasileiros e que esses polimorfismos podem ser utilizados como ferramentas para predição desses efeitos colaterais.

## INSTITUTO OSWALDO CRUZ

“Influência da variabilidade genética humana dos genes *NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes com tuberculose ativa”

### ABSTRACT

Tese de doutorado

Raquel Lima de F. Teixeira

Isoniazid (INH), one of the most important drugs used in anti-TB treatment, is also the major drug involved in hepatotoxicity. Genetic variability in several loci coding for drug-metabolizing enzymes can contribute to individual differences in drug response or toxicity and have been attributed to therapeutic failure and adverse drug reactions (ADRs). INH is oral administered and is metabolized by hepatic enzymes. The N-acetyltransferase 2 (*NAT2*) and the monooxygenase CYP450 2E1 (*CYP2E1*) are involved in the mechanism of hepatic injury induced by INH. *NAT2* is a polymorphic gene and mutations in its coding region can alter its enzymatic activity classifying individuals as slow, intermediate and fast acetylators. The slow acetylators can accumulate hepatotoxic intermediates developing liver disorders. Glutathione S-transferase (GST) is another enzyme supposed to play a protective role in INH biotransformation. Considering the differences found in the *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1* alleles' distribution among interethnic populations and their clinical importance in INH biotransformation, our goal was (i) to describe the genetic profile of the *NAT2* gene in individuals from two different regions of Brazil: Rio de Janeiro and Goiás States; (ii) to identify new *NAT2* variants in Brazilian individuals resident of two different regions by direct sequencing, and to characterize new haplotypes after cloning and re-sequencing as well as by molecular modeling; and (iii) to evaluate the contribution of genetic variants of *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1* in the susceptibility to anti-TB drug-induced hepatotoxicity in Brazilian individuals. Seventeen Single Nucleotide Polymorphisms (SNPs) already described were identified in Brazilian individuals in which seven (191G>A; 282C>T; 341T>C; 481C>T; 590G>A; 803A>G and 857G>A) were the most common SNPs in world population. Moreover, six new mutations were identified and seven new haplotypes were characterized in individuals resident in Rio de Janeiro and Goiás by cloning and re-sequencing and molecular modeling experimental. Our results showed the high incidence of *NAT2* alleles with low activity in Brazilian individuals and the variability of their frequencies between the two regions analyzed. Therefore, we could also observe that slow acetylation status was a risk factor for both anti-TB drug-induced hepatotoxicity (OR, 2.62; 95% IC, 1.75-3.49;  $p = 0.03$ ) and hepatitis (OR, 3.59; 95% IC, 2.53-4.64;  $p = 0.02$ ) when compared with intermediate/fast acetylation. No significant association between *CYP2E1*, *GSTM1* and *GSTT1* genotypes and liver side effects occurrence was observed. Even so, the results presented here support the influence of the genetic factors in the occurrence of antituberculosis drug-induced liver adverse reactions among Brazilians and can be used to predict these side effects.

*À minha avó Margarida, que não se encontra mais entre nós, mas que sempre curtiu e admirou meu sucesso profissional, e aos seus cinco irmãos que morreram de tuberculose, em uma época em que a cura ainda era um desafio – a eles dedico esta obra.*



## AGRADECIMENTOS

A Deus, por tudo que representa na minha vida, guiando-me e dando forças, iluminando minhas mãos para que eu pudesse concluir esse trabalho. Obrigada Senhor por tudo que me foi proporcionado nessa etapa da minha vida, vindo a contribuir para meu amadurecimento e aumentando ainda mais minha fé em Ti.

À minha mãe Cristina e meus irmãos Alexandre e Gabriel, que amo tanto, pela paciência, carinho e incentivo, durante todos esses anos dedicados à pesquisa, e a mais nova e querida integrante da família, minha cunhada Adriana. Dri, obrigada pela força. Você é MARA!

Aos meus orientadores “Adalba” e “Antony de Mira”, pela oportunidade da realização desse trabalho, pelas risadas, pelos chopps e, acima de tudo, por contribuírem de forma essencial para minha formação profissional. Obrigada pelos ensinamentos científicos e de vida! Hoje, eu perco dois orientadores, mas ganho dois grandes amigos!

Ao Dr. Pedro Cabello, meu chefe e amigo! Obrigada Pedro, por permitir que eu continuasse no doutorado e concluísse esse trabalho após ter assumido o cargo de Pesquisadora Assistente em seu laboratório. Obrigada pelas trocas diárias no convívio do laboratório e por sempre acreditar no meu potencial. A você toda a minha gratidão;

Ao meu grande Amor, José Carlos Xavier, por ser meu companheiro, parceiro, cúmplice; amante e amigo. Que essa felicidade que estamos vivendo agora permaneça em nossos corações pelo resto de nossas vidas! TE AMO!

À minha nova família: aos meus sogros, Fátima e José Carlos, aos meus novos “vós” Alda e José e aos meus cunhados queridos, Roberta e Carlos Henrique. Fátima, obrigada pelo carinho e pela correção da tese, de acordo com as novas regras da língua portuguesa;

À minha “filha-aluna” Renata Morato, por estar ao meu lado, sempre se esforçando para a conclusão dessa tese. Obrigada pela paciência e pelas incansáveis padronizações de PCRs, seqüenciamento, purificação, etc. O sucesso é nosso! Te adoro!

À Marcinha, minha mãe-amiga que, mesmo de longe, sempre esteve disponível quando precisei. Obrigada pelos ensinamentos, pelo carinho e pela amizade!

Às “amigas-da-carona” - Lizânia, Carol, Mary, Paulinha, Renata e Larissa - pelas conversas e desabafos nas nossas idas e vindas, Niterói/FIOCRUZ/Niterói, e principalmente pelas nossas saídas semanais de quinta que são sempre tão animadas! Carol, valeu pela formatação das referências!!! Sua ajuda foi essencial!

Às amigas do “Chopp das Amigas” - Irene, Marcinha, Elisa, Juliana e Júlia - que sempre me apoiaram e acreditaram no meu potencial; vocês são mais que amigas....são irmãs!!!

À minha psicóloga, Dra. Joana, que me ajudou nesses últimos cinco anos a ser uma pessoa melhor, me ensinando a ver a vida com um olhar diferenciado!!! E não é que funcionou?!

A todos do Laboratório de Genética Humana pelas risadas, pelas festinhas, pelas discussões científicas e pelas horas agradáveis durante o trabalho;

À Dra Fernanda Mello, por sempre estar disposta a ajudar e esclarecer questões da clínica médica que pertencem a um mundo bem diferente do meu. Obrigada por ter sido mais que colaboradora e por vibrar junto comigo as conquistas alcançadas durante a realização desse projeto;

Às enfermeiras Adriana Rezende e Lígia Mayumi: sem elas, a realização deste trabalho seria impossível! Meninas, obrigada pelo carinho, dedicação e paciência!

Ao Marcelo, ao Érico, à Aline e à Andreza – o *dream team* da Plataforma de Sequenciamento - por estarem sempre dispostos a ajudar no sequenciamento realizado durante toda a tese;

Ao Floriano e à Aline, do Laboratório de Bioquímica de Proteínas e Peptídeos, que realizaram a parte de modelagem molecular e análise estrutural de NAT2, de forma a enriquecer ainda mais o trabalho. Muito obrigada pela disponibilidade e por “vestirem a camisa” do projeto!

À Débora Goulart, minha professora de inglês e amiga pessoal. Obrigada pelas correções dos artigos e, acima de tudo, por transformar uma aula de inglês em algo tão prazeroso!

A todos do Laboratório de Biologia Molecular Aplicada a Micobactérias, por me acolherem sempre que precisei, principalmente durante o primeiro ano de doutorado, em especial ao querido amigo Harrison e ao meu ex-chefe Dr. Philip Noel Sulffys;

À Dra. Leila de Mendonça Lima, por ceder o laboratório para realização dos experimentos de clonagem;

Aos pacientes, contatos domiciliares e profissionais de saúde envolvidos neste estudo e que contribuíram para um melhor entendimento da farmacogenética na tuberculose;

Ao Instituto Oswaldo Cruz que me concedeu a bolsa nos primeiros meses de doutorado;

À Faperj, ao CNP-q e ao programa POM da Fiocruz pelo apoio financeiro.

*"É melhor tentar e falhar, que preocupar-se e ver a vida passar; é melhor tentar, ainda que em vão, que sentar-se fazendo nada até o final. Eu prefiro na chuva caminhar, que em dias tristes em casa me esconder. Prefiro ser feliz, embora louco, que em conformidade viver ..."*  
*Martin Luther King*

*"Na atividade da vida, o que importa não são o valor e a perfeição material das obras, mas a grandeza do amor com que as realizamos"*  
*Frei Anselmo*

## ABREVIATURAS

Acetil-CoA	Acetil Coenzima A
ACP	<i>Acyl Carrier Protein</i> – Proteína carreadora de grupamento acil
ADR	<i>Adverse Drug Reaction</i> – Reação adversa a droga
Asp	Aspartato ou Ácido aspártico
ATG	Códon de iniciação (adenosina, timina e guanina)
CD4 <sup>+</sup>	Receptor de superfície de células T4
CD8 <sup>+</sup>	Receptor de superfície de células T8
CYP 450	Citocromo P450
CYP2E1	Monooxigenase P450 da família 2, subfamília E, isoenzima 1
Cys	Cisteína
EUA	Estados Unidos da América
GSH	Glutationa
GST	Glutationa S transferase
GSTM1	Glutationa S transferase da família <i>mu</i> , isoenzima 1
GSTP1	Glutationa S transferase da família <i>pi</i> , isoenzima 1
GSTT1	Glutationa S transferase da família <i>theta</i> , isoenzima 1
His	Histidina
IC 95%	Intervalo de Confiança 95%
IFN- $\gamma$	Interferon gama
INH	Isoniazida
InhA	Enoil-ACP redutase
INPI	Instituto Nacional de Propriedade Industrial
KasA	$\beta$ -cetoacil-ACP sintetase
KatG	catalase-peroxidase micobacteriana
Kb	Kilo base
KDa	Kilo Dalton
MDR	Multi-droga resistente
Mn	Manganês
NAD <sup>+</sup>	Nicotinamida adenina dinucleotídeo oxidado
NADH	Nicotinamida adenina dinucleotídeo reduzido
NAT	N-acetiltransferase
NAT1	N-acetiltransferase 1
NAT2	N-acetiltransferase 2

OMS	Organização Mundial da Saúde
OR	<i>Odds Ratio</i> – Razão de Chance
<i>P</i>	Propabilidade
pb	Par de base
PCR	<i>Polymerase Chain Reaction</i> – Reação em cadeia da polimerase
PCT	<i>Patent Cooperation Treaty</i> – Tratado de Cooperação de Patente
<i>pNAT</i>	Pseudogene que constitui a família de genes <i>NATs</i>
PPD	<i>Purified Protein Derivative</i> - Derivado purificado da proteína
RFLP	<i>Restriction Fragment Length Polymorphism</i> – Polimorfismo de tamanho de fragmento por restrição
RIF	Rifampicina
RNI	<i>Reactive nitrogen intermediate</i> - Intermediários reativos de nitrogênio
ROI	<i>Reactive oxygen intermediate</i> - Intermediários reativos de oxigênio
SNP	Single Nucleotide Polymorphism – Polimorfismo de base única
TBMR	Tuberculose multi-resistente
TNF- $\alpha$	Fator de Necrose Tumoral alfa

## SUMÁRIO

	<b>Página</b>
Resumo.....	iv
Abstract.....	v
Abreviaturas.....	x
Lista de Figuras.....	xiii
Lista de Tabelas.....	xv
1) INTRODUÇÃO.....	1
1.1. Tuberculose: a doença.....	2
1.2. Tratamento da TB.....	5
1.3. Mecanismos de ação da isoniazida.....	7
1.4. Reações Adversas a Drogas: mecanismos da lesão hepatocelular mediada por drogas e o papel da isoniazida nesse contexto.....	10
1.5. A ciência da Farmacogenética.....	16
1.6. Genética molecular da N-acetiltransferase 2.....	20
1.7. Genética molecular da monoxigenase CYP2E1.....	28
1.8. Genética molecular das glutationas S-transferases.....	32
2) OBJETIVOS.....	36
3) RESULTADOS.....	39
3.1. Capítulo 1.....	40
3.2. Capítulo 2.....	51
3.3. Capítulo 3.....	90
3.4. Capítulo 4.....	137
4) DISCUSSÃO.....	165
5) CONCLUSÕES.....	175
6) REFERÊNCIAS BIBLIOGRÁFICAS.....	177
7) ANEXOS.....	193

## LISTA DE FIGURAS

	Página
1) Estimativa do número de novos casos de tuberculose no mundo em 2006.....	3
2) Incidência da tuberculose nos diferentes estados do Brasil em 2006.....	3
3) Eventos cronológicos após contato com <i>M. tuberculosis</i> .....	5
4) Estrutura da isoniazida.....	7
5) Mecanismo de ação da isoniazida em <i>M. tuberculosis</i> .....	9
6) Seis mecanismos de lesão no fígado causada por drogas.....	11
7) Reações de biotransformação de fármacos.....	13
8) Metabolismo da isoniazida no fígado.....	15
9) Áreas de conhecimento da farmacologia.....	17
10) Esquema mostrando as três possibilidades resultantes do metabolismo de fármacos administrados como droga ativa ou pró-droga.....	19
11) Esquema mostrando o benefício do ajuste de dose de uma droga baseado no genótipo do indivíduo.....	19
12) A prática da farmacogenética.....	20
13) Localização cromossômica dos <i>loci</i> das N-acetiltransferases humanas.....	21
14) Estrutura geral das N-acetiltransferases 1 e 2 humanas e de <i>S. typhimurium</i> .....	24
15) Nomenclatura das isoenzimas da super família CYP450.....	29
16) Abundância relativa das isoformas de CYP450 no fígado e alguns exemplos de substratos, inibidores e indutores.....	30
17) Localização dos polimorfismos no <i>locus</i> de <i>CYP2E1</i> e as respectivas enzimas de restrição utilizadas para detectar dos diferentes genótipos.....	31
18) Localização estrutural do <i>cluster</i> gênico de 100 kb que codifica para a subfamília GST mu.....	33
19) Localização estrutural do <i>cluster</i> gênico que codifica para a subfamília GST teta.....	34
20) Figura 1 do Capítulo 2: Sequência nucleotídica da região codificante de <i>NAT2</i> indicando as mutações novas protegidas pela patente BRPI0704093-8.....	86
21) Figura 1 do Capítulo 3: Caracterização dos haplótipos/alelos de <i>NAT2</i> contendo mutações novas ou raras em indivíduos brasileiros.....	126
22) Figura 2 do Capítulo 3: Mapeamento das mutações encontradas na região codificante de <i>NAT2</i> na estrutura tridimensional da proteína.....	127
23) Figura 3 do Capítulo 3: Alterações nas pontes de hidrogênio e no potencial	129

eletrostático frente à troca de aminoácido Tyr208His na estrutura da proteína NAT2 humana.....	
24) Figura 4 do Capítulo 3: Modelo estrutural da variante Gly51Val em NAT2 humana.....	130
25) Figura 5 do Capítulo 3: Modelo estrutural da variante Cys68Tyr em NAT2 humana.....	131
26) Figura do Material Suplementar do Capítulo 3: Alinhamento múltiplo de várias sequências de <i>NAT</i> .....	132
27) Estratégia de amplificação e seqüenciamento de <i>NAT2</i> (Anexo 2).....	197
28) Eletroforese em gel de agarose 1% contendo os produtos de PCR <i>NAT2</i> purificados (Anexo 2).....	197
29) Ilustração da utilização do <i>software</i> SeqScape v. 2.5. (Anexo 2).....	198
30) Identificação de SNP através da visualização do cromatograma dentro do programa SeqScape v. 2.5. (Anexo 2).....	198
31) Sequência nucleotídica da região de clonagem do vetor pCR-BluntII-TOPO (Anexo 4).....	202
32) Gel de agarose 1,5% com PCR <i>NAT2</i> a partir de colônia para identificação de bactérias portadoras de pCRBlunt-II/ <i>NAT2</i> (Anexo 4).....	202
33) Análise de restrição para confirmação da presença do inserto dos clones pCRBlunt-II/ <i>NAT2</i> (Anexo 4).....	202
34) Gel de agarose 2% com o perfil de digestão ( <i>Rsa</i> I) dos três genótipos de <i>CYP2E1</i> para o SNP -1053 C→T (Anexo 6).....	207
35) Gel de agarose 2,5% com PCR padrão para detectar a presença do gene <i>GSTM1</i> tipo-selvagem (Anexo 6).....	207
36) Gel de agarose 1,5% com PCR <i>Multiplex</i> para identificação dos genótipos de <i>GSTT1</i> (Anexo 6).....	207



## LISTA DE TABELAS

	Página
1) Esquemas de tratamento padronizados para tuberculose no Brasil.....	6
2) Características moleculares das N-acetiltransferases humanas e relação entre fenótipo e genótipo apresentados por essas enzimas.....	22
3) Alelos de <i>NAT2</i> já descritos na literatura.....	25
4) Comparação interétnica da distribuição dos alelos de <i>NAT2</i> e a frequência do fenótipo de acetilação lenta encontrada nas diferentes populações.....	27
5) Tabela 1 do Capítulo 1: Frequência das mutações no gene da N-acetiltransferase 2 em indivíduos brasileiros provenientes do Rio de Janeiro e Goiás.....	44
6) Tabela 2 do Capítulo 1: Frequência das variantes alélicas de <i>NAT2</i> entre os indivíduos do Rio de Janeiro e Goiás.....	45
7) Tabela 3 do Capítulo 1: caracterização genotípica do gene <i>NAT2</i> a determinação preliminar do <i>status</i> de acetilação em 202 indivíduos de duas regiões diferentes do Brasil.....	46
8) Tabela 4 do Capítulo 1: Frequência alélica das sete mutações mais comuns do gene <i>NAT2</i> em diferentes populações.....	47
9) Tabela 1 do Capítulo 2: Sequências dos oligonucleotídeos utilizados para a amplificação e/ou sequenciamento da região codificante do gene <i>NAT2</i> .....	79
10) Tabela 2 do Capítulo 2: Lista dos polimorfismos de base única (SNPs) descobertos no gene <i>NAT2</i> .....	81
11) Tabela 3 do Capítulo 2: Frequência alélica dos SNPs da presente invenção no gene <i>NAT2</i> em duas diferentes regiões do Brasil.....	84
12) Tabela 4 do Capítulo 2: Frequência dos SNPs identificados no gene <i>NAT2</i> em duas diferentes regiões do Brasil.....	84
13) Tabela 1 do Capítulo 3: Todas as mutações já identificadas na região codificante de <i>NAT2</i> e suas frequências em 404 indivíduos brasileiros residentes no Rio de Janeiro e em Goiás.....	136
14) Tabela 1 do Capítulo 4: Características clínicas dos pacientes com TB com e sem hepatotoxicidade. Frequências alélicas e genotípicas dos polimorfismos nos genes <i>CYP2E1</i> , <i>GSTT1</i> e <i>GSTM1</i> nos dois grupos.....	158
15) Tabela 2 do Capítulo 4: Características clínicas dos pacientes com TB com e sem hepatite medicamentosa. Frequências alélicas e genotípicas dos polimorfismos nos genes <i>CYP2E1</i> , <i>GSTT1</i> e <i>GSTM1</i> nos dois grupos.....	159

16) Tabela 3 do Capítulo 4: Frequência dos alelos de <i>NAT2</i> nos indivíduos brasileiros arrolados no estudo.....	160
17) Tabela 4 do Capítulo 4: Genótipo/fenótipo de <i>NAT2</i> e suscetibilidade a hepatotoxicidade induzida por isoniazida.....	161
18) Tabela 5 do Capítulo 4: Genótipo/fenótipo de <i>NAT2</i> e suscetibilidade a hepatite medicamentosa induzida por isoniazida.....	163



## *1- INTRODUÇÃO*



## 1.1) Tuberculose: a doença

A tuberculose (TB) é uma doença infectocontagiosa causada por uma bactéria aeróbia patogênica, que estabelece a sua infecção preferencialmente nos pulmões, sendo considerada uma das mais importantes doenças infecciosas no mundo. É estimado que um terço da população mundial esteja infectado e apresente tuberculose latente (WHO, 2000). Segundo estimativas da Organização Mundial de Saúde (OMS), apesar de existir um tratamento com elevada eficácia, entre 2000 e 2020, um bilhão de pessoas estarão infectadas pelo bacilo da tuberculose, cerca de duzentos milhões desenvolverão a doença e trinta e cinco milhões irão a óbito caso não sejam identificados novos métodos diagnósticos, novas vacinas, novos medicamentos e que políticas mais apropriadas para controle dos pacientes com tuberculose ativa e seus contatos sejam desenvolvidas (WHO, 2000).

De acordo com a OMS, foram notificados nove milhões e duzentos mil novos casos de TB no mundo em 2006 (Figura 1) com um milhão e setecentos mil óbitos, principal fator que torna a TB um problema de saúde pública mundial. O continente africano contém as maiores taxas da doença, apresentando trezentos e sessenta e três casos por cem mil habitantes. No total, 83% dos casos novos de tuberculose são atribuídos às regiões da África e Ásia, onde a Índia, China, Indonésia, Nigéria e África do Sul ocupam os cinco primeiros lugares em termos de casos absolutos de tuberculose (WHO, 2008). A tuberculose pode ser prevenida e a maioria dos casos tem cura; entretanto, fatores demográficos e sócio-econômicos dificultam sua prevenção e tratamento. Além disso, a disseminação da epidemia do vírus HIV e a emergência de cepas multidroga-resistentes (MDRs) em várias partes do mundo têm contribuído para o aumento do número de casos de tuberculose (Schluger, 2001; van Crevel, *et al.*, 2002; Tufariello, Chan & Flynn, 2003).

O Brasil ocupa hoje o décimo-sexto lugar na lista dos vinte e dois países com a mais elevada taxa de incidência de tuberculose. Estima-se uma prevalência de cinquenta milhões de infectados, com cerca de cento e dez mil casos novos e seis mil óbitos ocorrendo anualmente (Santos-Filho, 2006; WHO, 2008). Os estados mais afetados incluem Amazonas (73,5 casos/100.000 habitantes) o Rio de Janeiro (73,2 casos/100.000 habitantes) e o Ceará (50,8 casos/100.000 habitantes), e aqueles que apresentam as menores taxas de incidência de tuberculose são Tocantins (18,1 casos/100.000 habitantes), Distrito Federal (16,7 casos/100.000 habitantes) e Goiás (15,3 casos/100.000 habitantes) (SINAN - Sistema de Informação de Agravos de Notificação - Figura 2).

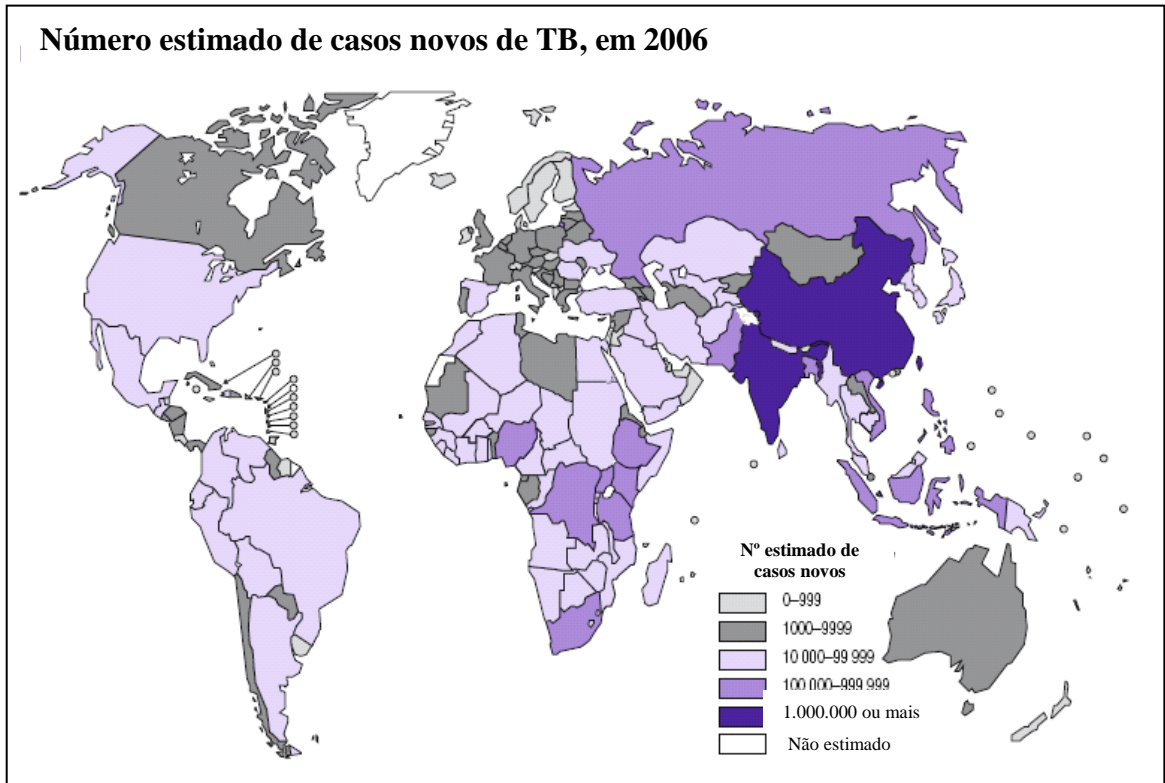


Figura 1: Estimativa do número de novos casos de tuberculose no mundo, em 2006 (Controle Global da TB, WHO report 2008)

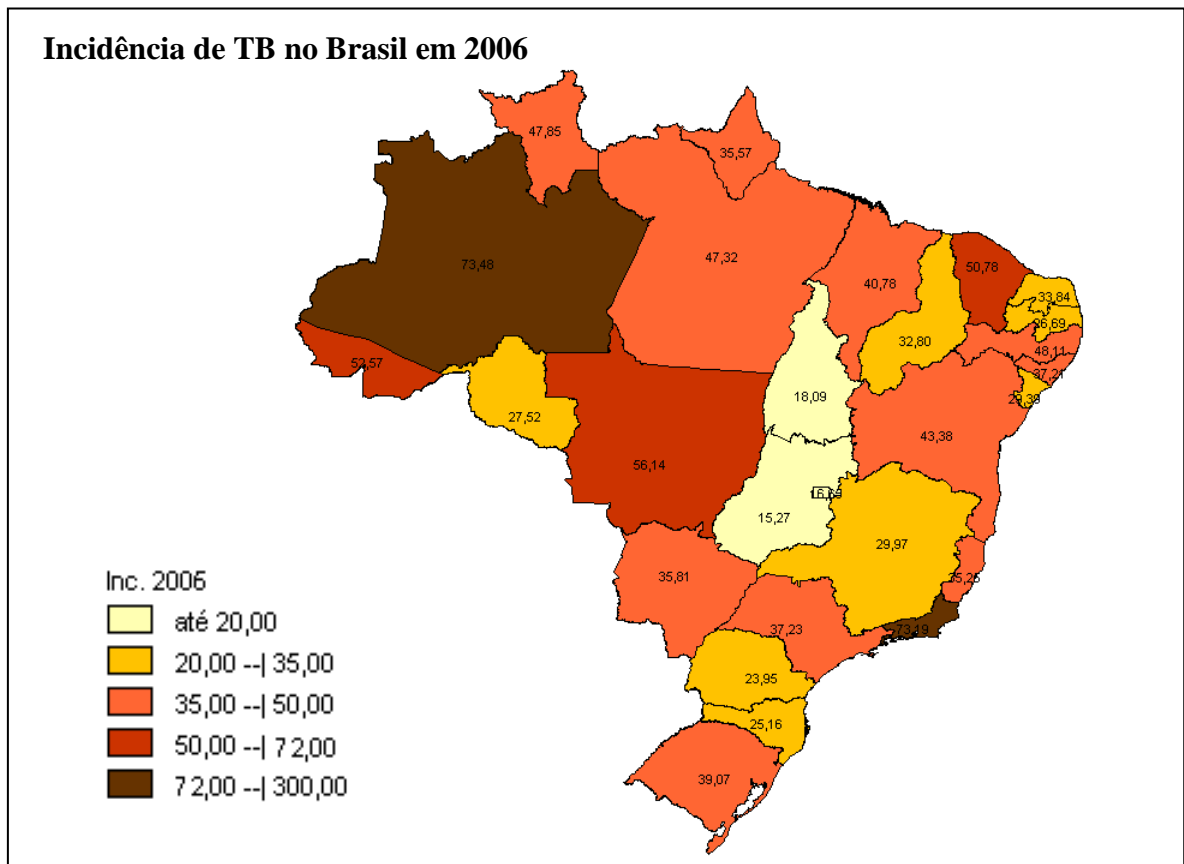


Figura 2: Incidência de tuberculose nos diferentes estados do Brasil em 2006. Os números representam os casos de TB por 100.000 habitantes. Fonte: SINAN (Sistema de Informação de Agravos de Notificação).



A tuberculose é causada pelo *Mycobacterium tuberculosis*, um bacilo álcool-ácido resistente, de crescimento lento, que é transmitido principalmente através das vias respiratórias superiores. A infecção acontece inicialmente nos macrófagos alveolares, após inalação da bactéria contida em perdigotos liberados por pacientes com tuberculose ativa não tratada. Contudo, o micro-organismo pode atingir qualquer órgão através da disseminação hematogênica e quando a TB é estabelecida, a localização, a gravidade e a evolução da doença são altamente variáveis.

O *M. tuberculosis* se replica no interior dos macrófagos, induzindo a produção de citocinas e quimiocinas que iniciam uma resposta inflamatória nos pulmões. Macrófagos, linfócitos, monócitos, bem como outras células inflamatórias, migram para o local da infecção e formam o granuloma, cuja principal função é segregar a infecção e prevenir a disseminação do bacilo para outros órgãos, assim como concentrar a resposta imune diretamente no sítio da infecção. Neste contexto, a principal citocina envolvida na manutenção da estrutura do granuloma e na ativação de macrófagos é o Fator de Necrose Tumoral alfa (TNF- $\alpha$ ), liberado por fagócitos ativados pelo *M. tuberculosis* (van Crevel, *et al.*, 2002).

As células T são fundamentais na resposta imune contra o *M. tuberculosis*. Os linfócitos T CD4<sup>+</sup> exercem seu efeito protetor através da produção de citocinas, principalmente Interferon gama (IFN- $\gamma$ ), após a estimulação com antígenos micobacterianos. Já os linfócitos T CD8<sup>+</sup>, além de secretarem citocinas como IFN- $\gamma$  e TNF- $\alpha$ , desempenham uma atividade citotóxica, levando à lise de células infectadas ou à morte intracelular do patógeno, diretamente (Schluger, 2001).

A citocina proinflamatória IFN- $\gamma$  desempenha um papel chave na proteção contra a tuberculose. Secretado por diversas células imunológicas, o IFN- $\gamma$  estimula a produção de intermediários reativos de oxigênio (ROI) e nitrogênio (RNI), através da indução da expressão da óxido nítrico sintase (NOS2) por macrófagos. Esses reativos são importantes para a realização das funções efetoras de inibição do crescimento intracelular e morte da micobactéria pelo macrófago. Adicionalmente, o IFN- $\gamma$  induz um aumento da expressão de moléculas do complexo MHC de classe II pelos macrófagos e também regula positivamente a expressão e secreção de TNF- $\alpha$  (Schluger, 2001; Flynn & Chan, 2001; Tufariello *et al.*, 2003).

Existem basicamente três possíveis desfechos após o contato da pessoa com o *M. tuberculosis* (Figura 3): (i) o micro-organismo pode ser imediatamente destruído pela resposta imune inata do hospedeiro; (ii) uma parcela das pessoas infectadas (cerca de 10%) desenvolve a tuberculose ativa (TB primária) no intervalo de um a três anos (esse grupo, aparentemente,



não é capaz de controlar a infecção inicial nem de desenvolver uma imunidade protetora a tempo de prevenir a doença ativa); e finalmente (iii) a maioria das pessoas infectadas com *M. tuberculosis* apresentam uma infecção latente. Aproximadamente de 5 a 10 % dessas pessoas irão sofrer uma reativação da infecção e apresentar tuberculose ativa (TB pós-primária) que normalmente acontece em resposta a perturbações do sistema imune (Flynn & Chan, 2001a; Flynn & Chan, 2001b; van Crevel *et al.*, 2002).

O PPD (*Purified Protein Derivative* - Derivado Purificado da Proteína) ou Prova Tuberculínica Cutânea é um método auxiliar para diagnóstico de TB. Baseia-se na reação celular desenvolvida após a inoculação intradérmica de um derivado protéico de *M. tuberculosis*. Essa reação cutânea é registrada em milímetros, para a posterior interpretação clínica. O resultado positivo caracteriza apenas infecção pelo bacilo (Dalcomo *et al.*, 2004).

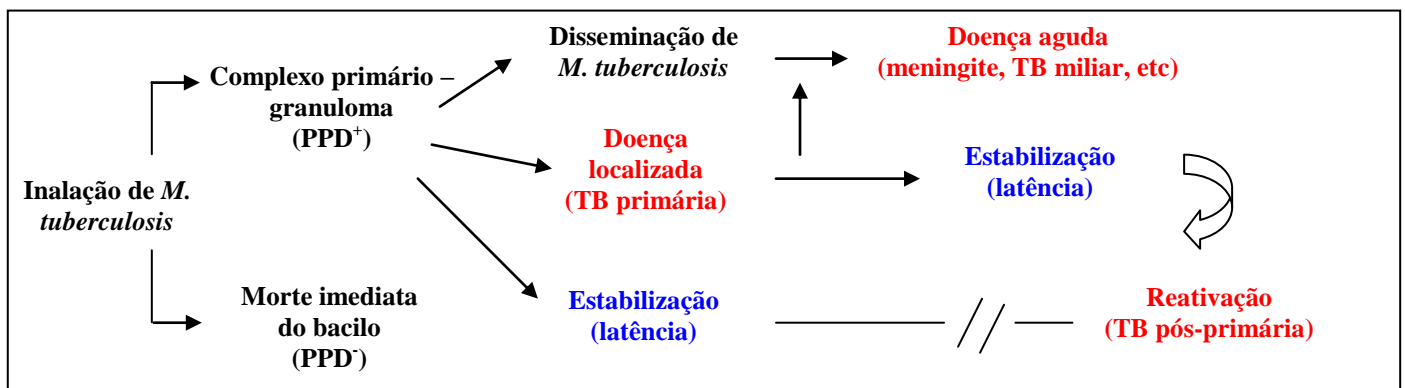


Figura 3: Eventos cronológicos após contato com *M. tuberculosis*. Figura adaptada de van Crevel, Ottenhoff e van der Meer, 2002.

PPD – Prova tuberculínica cutânea

## 1.2) Tratamento da TB

Desde 1979, os esquemas de tratamento da TB preconizados pelo Ministério da Saúde no Brasil são: um, de primeira linha, para os casos virgens de tratamento, que compreende a administração de rifampicina (R), isoniazida (H) e pirazinamida (Z) por dois meses, seguidos de rifampicina e isoniazida por mais quatro meses (Esquema I – 2RHZ/4RH); e outro, de segunda linha, com administração de estreptomicina (S), pirazinamida, etambutol (E) e etionamina (Et) por três meses e etambutol e etionamida mais nove meses (Esquema III – 3SZEet/EEt), indicado para pacientes com falência ao Esquema I. Para os casos de meningoencefalite isolada ou associada a outras formas, é proposto um esquema especial (Esquema II) que compreende a administração de R, H e Z por dois meses, seguidos da administração de R e H por mais sete meses. Para os pacientes que não obtêm cura com os esquemas tradicionais descritos acima, chamados de portadores de TBMR (TB multi-



resistente), foi proposto, após ter sido validado pelo Ministério da Saúde, em 2000, um esquema alternativo com associação de amicacina (AM), ofloxacina (OFX), terizidon (TRZ), etambutol (E) e clofazimina (CLZ), com duração de 18 meses (Tabela 1 - Dalcomo *et al.*, 2004).

Tabela 1: Esquemas de tratamento para a tuberculose no Brasil (Dalcomo *et al.*, 2004).

Esquemas	Finalidade
Esquema 1 RHZ (2RHZ/4RH)	Diário, por seis meses para todas as formas pulmonares e extrapulmonares, todas as idades, e na coinfeção com HIV/AIDS
Esquema 2 RHZ (2RHZ/7RH)	Diário, por nove meses para meningoencefalite tuberculosa
Esquema 1R RHZE (2RHZE/4RHE)	Diário, por seis meses para retratamento por recidiva após a cura ou o reinício do tratamento por abandono
Esquema 3 SEEtZ (3EEtZ/9EEt)	Doze meses de tratamento, para falência bacteriológica aos esquemas anteriores
SEO (3SEOFX/9EOFX)	Doze meses de tratamento, para hepatopatias ou intolerância severa
SHE (3SHE/9HE)	Doze meses de tratamento, para hepatopatias
SRE (2SRE/4RE)	Seis meses de tratamento, para hepatopatias
12AM / 18OFX / 18 TRZ / 18 E / 18 CLZ	18 meses de tratamento, para TBMR

<sup>R</sup> – Rifampicina  
<sup>H</sup> – Isoniazida  
<sup>Z</sup> – Pirazinamida  
<sup>E</sup> – Etambutol

<sup>S</sup> – Estreptomicina  
<sup>O, OFX</sup> – Ofloxacina  
<sup>AM</sup> – Amicacina  
<sup>TRZ</sup> – Terizidon

<sup>CLZ</sup> – Clofazimina

A quimioprofilaxia da tuberculose é uma medida terapêutica que consiste no uso de medicamentos para a prevenção da infecção pelo *M. tuberculosis* ou para evitar o desenvolvimento da doença naqueles infectados. Baseia-se na administração de isoniazida (INH), durante seis meses, em pessoas não infectadas para prevenir a infecção (quimioterapia primária) ou em pessoas já infectadas (que apresentarem viragem do PPD), mas sem sinais de doença, para prevenir a evolução da infecção à mesma (quimioprofilaxia secundária) (Dalcomo *et al.*, 2004).





Os grupos de risco onde a quimioprofilaxia apresenta algum benefício incluem: pessoas infectadas com HIV com PPD positivo (endurecimento  $\geq 5$ mm); contatos de pessoas com TB ativa, sem sinais da doença ativa e fortes reatores à prova tuberculínica ( $\geq 10$  mm); crianças com menos de 15 anos não vacinadas; crianças vacinadas com BCG nos últimos 3 anos e com resposta tuberculínica igual ou superior a 15 mm de endurecimento; pessoas com viragem tuberculínica recente e que apresentam aumento na resposta à prova tuberculínica de no mínimo 10 mm em relação à endureção inicial. No Brasil, em particular, com a recomendação de não se vacinar os profissionais de saúde não reatores, torna-se necessário realizar inquéritos tuberculínicos em unidades de saúde e indicar a quimioprofilaxia aos profissionais de saúde que apresentarem viragem tuberculínica (Frieden *et al.*, 2003; Dalcomo *et al.*, 2004).

### 1.3) Mecanismos de ação da isoniazida

Em 1952, a isoniazida foi introduzida ao esquema quimioterápico de controle da tuberculose e até hoje apresenta uma excelente eficácia contra *M. tuberculosis*, tanto no tratamento da infecção ativa, quanto na profilaxia.

A isoniazida (hidrazida do ácido isonicotínico – Figura 4) possui ação bacteriostática para o bacilo em repouso e ação bactericida para o bacilo em multiplicação. A INH penetra com facilidade nas células e se mostra eficaz contra bacilos que crescem no interior de macrófagos, ou *in vitro* (meio de cultura). Além disso, esse fármaco é seletivo para micobactérias, sendo necessárias altas concentrações para que algum efeito seja observado em outros micro-organismos (Hardman *et al.*, 1996).

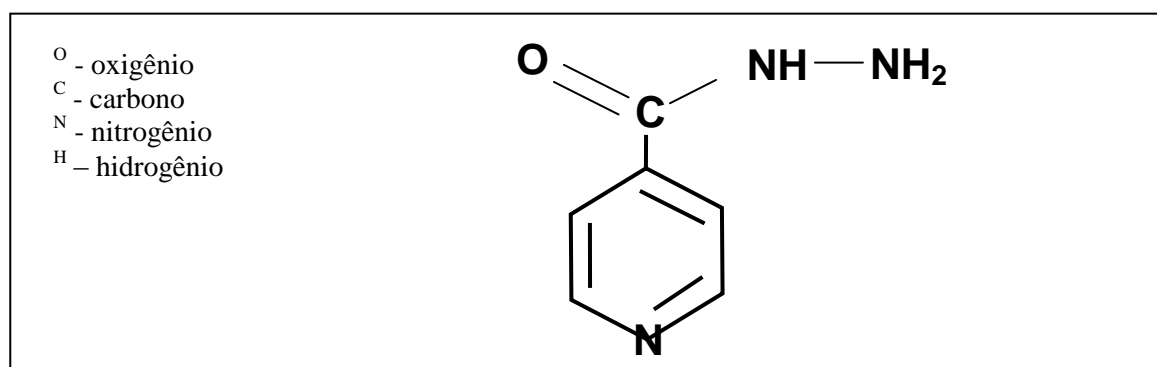


Figura 4: Estrutura da isoniazida.

Em 1970, foi demonstrado que a INH inibe a síntese dos ácidos micólicos em *M. tuberculosis* levando então à correlação entre a biossíntese de ácidos micólicos e à viabilidade celular (Winder & Collins, 1970; Winder *et al.*, 1970; Takayama *et al.*, 1972).

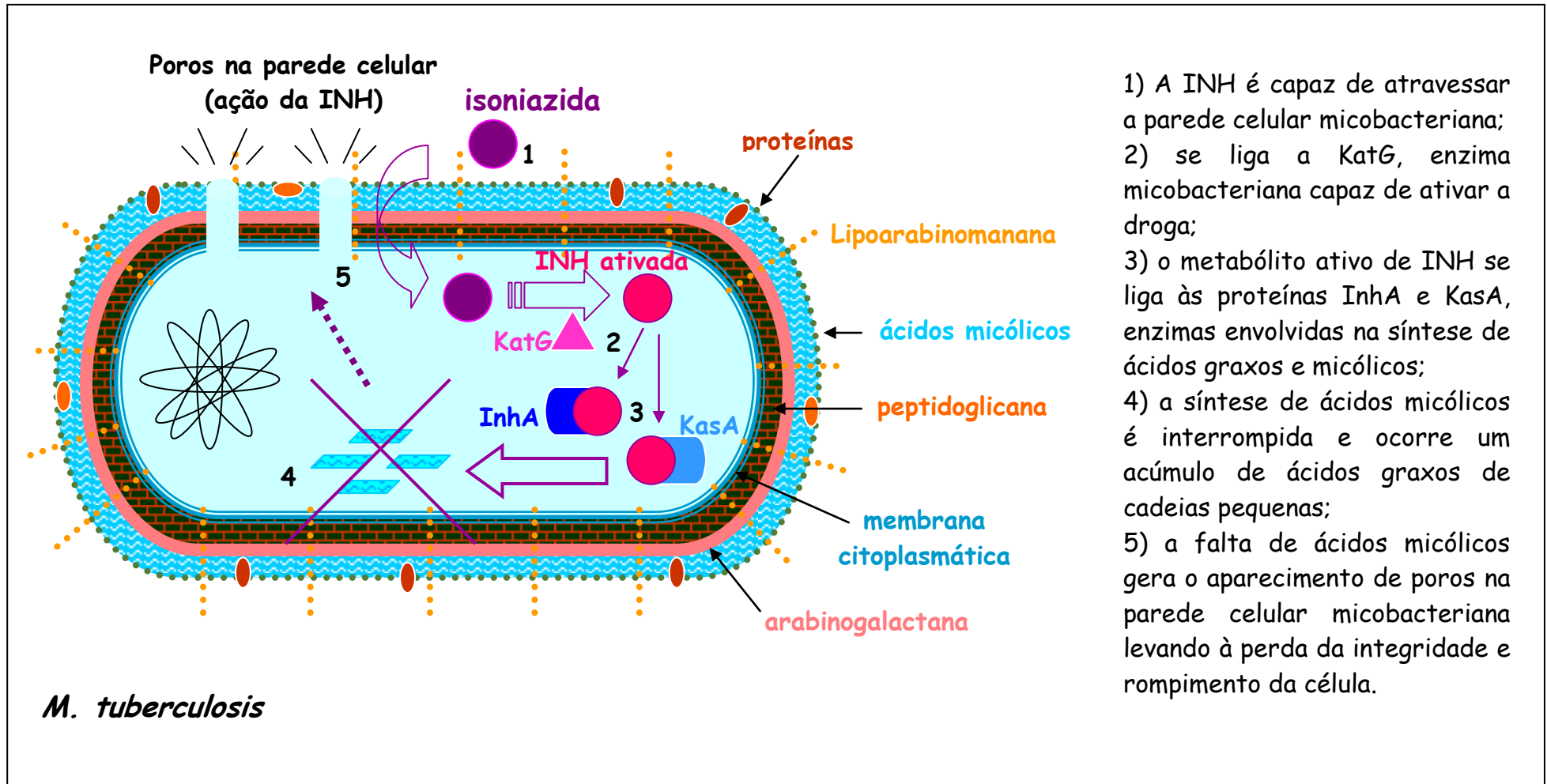


Os ácidos micólicos são componentes estruturais importantes para a manutenção da integridade da parede celular micobacteriana e consistem em cadeias longas de ácidos graxos ( $\alpha$ -pregueadas e  $\beta$ -hidroxiladas) que ligam o esqueleto rígido da parede celular (arabinogalactana) à membrana celular externa. Sendo assim, a inibição da síntese de ácidos micólicos leva a deformidades celulares em *M. tuberculosis*, causando a perda da integridade celular (Middlebrook, 1952). Conseqüentemente, a inibição da síntese do ácido micólico pode explicar a perda da álcool-ácido resistência pelas micobactérias expostas à isoniazida (Blanchard, 1996).

Outros mecanismos de ação da isoniazida, ainda não muito bem esclarecidos, incluem: depleção do cofator  $\text{NAD}^+$ , efeito na biossíntese de ácidos nucleicos, inibição da atividade de catalase-peroxidase micobacteriana, suscetibilidade aos efeitos tóxicos dos reativos de oxigênio (principalmente  $\text{H}_2\text{O}_2$ ) no fagolisossomo do macrófago (Zhang & Young, 1993; Rastogi & David, 1993; Blanchard, 1996 Hardman *et al.*, 1996).

A isoniazida é considerada uma pró-droga porque requer a ativação celular antes de exercer seus efeitos tóxicos no bacilo. A enzima micobacteriana catalase-peroxidase KatG (80KDa) transforma a INH na sua forma ativada através de uma reação de oxidação (Slayden & Barry, 2000). Essa enzima desempenha diversas funções dentro da micobactéria, como: atividades de catalase-peroxidase, de peroxidase dependente de  $\text{Mn}^{2+}$ , de peroxinitrilase e de mono-oxigenase (P450-“like”), sendo importante no estresse oxidativo e na virulência.

Segundo dados da literatura, a forma ativa proposta da INH (radical isonicotinil ou acilisonicotínico) é um metabólito intermediário da formação do ácido isonicotínico, da isonicotinamida e piridina-4-carboxialdeído, produtos da oxidação de INH por KatG (Rozwarski *et al.*, 1998; Slayden & Barry, 2000). Uma vez formado no interior da célula micobacteriana, o radical isonicotinil se liga ao cofator  $\text{NAD}^+$  (nicotinamida adenina dinucleotídeo oxidado) localizado no sítio ativo da proteína InhA, uma enoil-ACP (*acyl carrier protein*) redutase. Além disso, a forma ativada de INH se liga covalentemente a outra proteína chamada KasA, uma  $\beta$ -cetoacil-ACP sintetase. Ambas as proteínas estão envolvidas na elongação de ácidos graxos e biossíntese de ácidos micólicos. A ligação do complexo NADH-isonicotinil a InhA e KasA reduz as atividades de redutase e sintetase das enzimas, interferindo com a biossíntese de ácidos graxos e micólicos (Figura 5), provocando então o acúmulo de ácidos graxos saturados de cadeias não muito longas (Rozwarski *et al.*, 1998). A interrupção da síntese de ácidos micólicos na micobactéria em crescimento leva à formação de poros na parede celular, com a conseqüente lise da célula (Figura 5).



- 1) A INH é capaz de atravessar a parede celular micobacteriana;
- 2) se liga a *KatG*, enzima micobacteriana capaz de ativar a droga;
- 3) o metabólito ativo de INH se liga às proteínas *InhA* e *KasA*, enzimas envolvidas na síntese de ácidos graxos e micólicos;
- 4) a síntese de ácidos micólicos é interrompida e ocorre um acúmulo de ácidos graxos de cadeias pequenas;
- 5) a falta de ácidos micólicos gera o aparecimento de poros na parede celular micobacteriana levando à perda da integridade e rompimento da célula.

Figura 5: Mecanismo de ação da isoniazida em *M. tuberculosis*.



#### **1.4) Reações Adversas a Drogas: mecanismos da lesão hepatocelular mediada por drogas, e o papel da isoniazida nesse contexto**

As reações adversas à isoniazida incluem febre, erupções cutâneas, hepatite e exantemas morbiliformes, macupapulares, purpúricos e urticariformes. Além disso, podem ocorrer reações hematológicas. Somente no início da década de setenta ficou evidente a possível ocorrência de lesão hepática grave, resultando na morte de alguns pacientes que estavam tomando INH (Garibaldi *et al.*, 1972). Estudos adicionais em adultos e crianças confirmaram esta observação, mostrando que o processo patológico característico consiste em necrose em ponte e multilobular. Os pacientes que recebem isoniazida devem ser cuidadosamente avaliados a intervalos mensais para sintomas de hepatite (como anorexia, mal-estar, fadiga, náusea e icterícia) e alertados sobre a necessidade de interromper o fármaco se esses sintomas ocorrerem (Hardman *et al.*, 1996).

As Reações Adversas a Drogas (*adverse drug reactions* - ADRs) podem levar a hospitalização e apresentam um alto custo financeiro para a sociedade. Esses custos incluem despesas para o tratamento das ADRs e despesas para evitar esses efeitos colaterais durante a administração do fármaco (Lundkvist & Jonsson, 2004). A ocorrência de ADRs graves e fatais tem sido extensivamente analisada em pacientes hospitalizados. Uma meta-análise de aproximadamente quarenta estudos prospectivos em hospitais dos Estados Unidos da América (EUA) sugere que de 6-7% dos pacientes internados sofrem de ADRs graves e que 0,32% dos pacientes apresentam ADRs fatais (Hug *et al.*, 2003). Isso resulta em aproximadamente 100.000 mortes por ano nos EUA, e um custo anual de mais de cem bilhões de dólares para a sociedade norte-americana, devido a um tempo maior de hospitalização e redução na produtividade (Ingelman-Sundberg, 2001; Hug *et al.*, 2003). Além disso, foi estimado que as ADRs são responsáveis por mais de 7% das internações nos hospitais do Reino Unido e 13% das admissões em clínicas médicas na Suécia (Ingelman-Sundberg, 2001), mostrando que as reações adversas constituem um problema ainda maior do que se pensava, no contexto de quimioterapia e desenvolvimento de novas drogas.

Na França, um estudo realizado durante dez anos, na Unidade Hepática do *Hôpital Beaujon* (Paris), mostrou que dentre todos os pacientes internados com hepatite aguda, 10% eram casos de reação adversa a drogas. A prevalência de hepatotoxicidade medicamentosa em pacientes com mais de cinquenta anos foi superior a 40%. No Japão, e em outros países orientais, os fármacos são responsáveis por cerca de 10-20% dos casos de hepatites fulminantes (Larrey, 2002).



A lesão hepática é o tipo mais frequente de ADRs e pode levar a retirada de um fármaco do mercado. A doença hepática induzida por drogas é uma das principais complicações durante a quimioterapia, uma vez que o fígado é o órgão central na biotransformação e excreção da maioria das drogas e substâncias xenobióticas (Bagheri *et al.*, 2000; Bissell *et al.*, 2001; Larrey, 2002; Lee, 2003). Existem basicamente seis mecanismos que envolvem primariamente a lesão do hepatócito (Figura 6). As reações da mono-oxigenase citocromo P450 (CYP450) com determinados fármacos geram metabólitos tóxicos que podem se ligar a proteínas intracelulares, levando a uma disfunção da homeostase da bomba de cálcio com consequente desestruturação das fibras de actina na superfície do hepatócito e lise celular (Figura 6A). Algumas drogas podem afetar proteínas transportadoras na membrana canicular interrompendo o fluxo da bile e então causando a colestase. (Figura 6B).

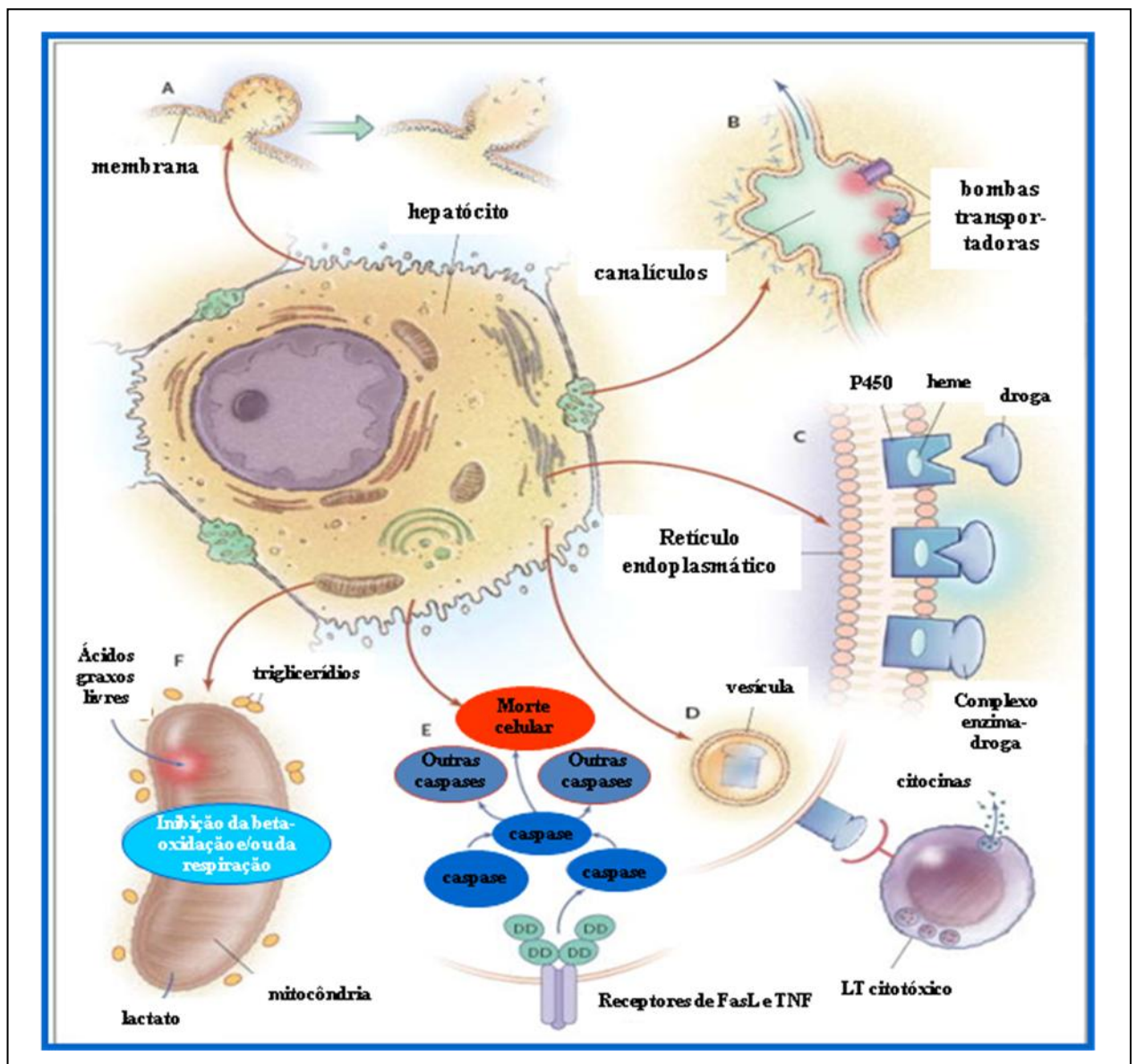


Figura 6: Seis mecanismos de lesão no fígado causada por drogas. Esquema retirado a partir de Lee, 2003.

DD – death domain



Várias reações envolvendo CYP P450 podem promover a ligação da droga à enzima, com consequente exposição desse complexo na superfície celular, reconhecimento por células T e produção de anticorpos (lesão por resposta autoimune) (Figura 6C/D). Por fim, determinadas drogas podem promover lesão hepática mediada por morte celular programada (apoptose) (Figura 6E) ou ainda serem capazes de inibir a respiração e/ou betaoxidação mitocondrial (Figura 6F).

As substâncias xenobióticas são, em sua maioria, lipofílicas. Essa propriedade facilita o seu transporte na corrente sanguínea por lipoproteínas e sua penetração nos órgãos, através das membranas lipídicas. Entretanto, as propriedades físico-químicas das moléculas dos fármacos dificultam sua posterior excreção do organismo. Sendo assim, essas substâncias são metabolizadas a compostos hidrossolúveis, apropriados à excreção renal ou biliar (Cotran *et al.*, 1996; Hardman *et al.*, 1996). As substâncias químicas são frequentemente metabolizadas, quase sempre através de múltiplas vias, a produtos que podem ser mais ou menos tóxicos do que o produto original. Um ou mais desses metabólitos podem, portanto interagir com macromoléculas-alvo (DNA, RNA, proteínas, receptores), resultando em efeito tóxico. O local da toxicidade é geralmente aquele onde ocorre a metabolização ou excreção dos metabólitos (Cotran *et al.*, 1996).

Os sistemas enzimáticos responsáveis pela biotransformação de muitos fármacos estão localizados no retículo endoplasmático liso do fígado (fração microssômica). Tais enzimas também são encontradas em outros órgãos como rins, pulmões e epitélio gastrintestinal, embora em menor concentração (Hardman *et al.*, 1996).

As reações químicas da biotransformação enzimática são classificadas como reações de Fase I ou reações de Fase II (Figura 7). As reações de Fase I convertem o fármaco original em um metabólito mais polar, através de oxidação, redução ou hidrólise. O metabólito resultante seria farmacologicamente inativo, menos ativo ou, às vezes, mais ativo que a molécula original (no caso de pró-droga). Esse metabólito pode ser eliminado ou participar das reações de Fase II. As reações de Fase II, denominadas reações de conjugação ou síntese, envolvem a ligação do fármaco ou metabólito primário a um substrato endógeno, como glicuronato, sulfato, acetato, aminoácidos ou glutadiona (tripeptídeo). Essas reações incluem reações de glicuronidação, metilação, sulfatação, acetilação, conjugação com glutadiona, conjugação com glicina (Hardman *et al.*, 1996).

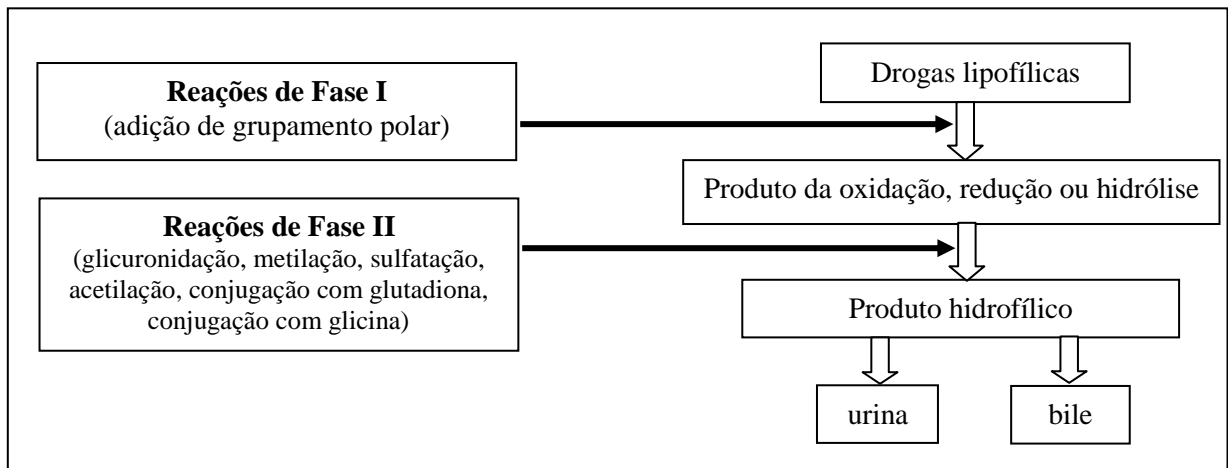


Figura 7: Reações de biotransformação de fármacos.

A lesão hepática pode ocorrer após a inalação, ingestão ou administração parenteral da droga. Em geral, dois tipos principais de hepatotoxicidade química são reconhecidos: o tóxico direto e o idiossincrásico (Harrison, 2002). A hepatotoxicidade direta ocorre com frequência de forma previsível em indivíduos expostos ao agente agressor, sendo dependente da dose. O período latente entre a exposição e a lesão hepática geralmente é curto (desde algumas horas até 48 horas). Os agentes que produzem hepatite tóxica são venenos sistêmicos ou são convertidos no fígado em metabólitos tóxicos. A hepatotoxicidade idiossincrásica é caracterizada por ser uma reação imprevisível, anormal e geneticamente determinada. A resposta observada é qualitativamente similar em todos os indivíduos e não é dose dependente. Considera-se que a maioria das reações idiossincrásicas resulta de diferenças na reatividade metabólica a agentes específicos. A suscetibilidade do hospedeiro é mediada pelas cinéticas de geração do metabólito tóxico, que difere entre os indivíduos. A formação de metabólitos reativos derivados de reações de biotransformação está frequentemente relacionada à hepatotoxicidade idiossincrásica, como é o caso da isoniazida. Além disso, nessas reações pode acontecer hipersensibilidade derivada da produção de auto-anticorpos; sendo assim, as reações idiossincrásicas podem ainda ser divididas em alérgicas e metabólicas (Harrison, 2002; Lee, 2003; Hug *et al.*, 2003).

O risco de hepatotoxicidade é influenciado por diferentes fatores adquiridos e genéticos. Dentre os fatores adquiridos podemos citar: idade, sexo, qualidade nutricional do indivíduo, interações medicamentosas, gravidez, doenças extra-hepáticas. Variações genéticas em isoenzimas que geram os bioprodutos tóxicos podem resultar em diferenças na reatividade metabólica a agentes específicos que variam de indivíduo para indivíduo (Larrey, 2000; Larrey, 2002; Lee, 2003). No caso da isoniazida, a idade é um fator de risco para a geração de hepatotoxicidade, principalmente em pessoas acima de 50 anos, enquanto que deficiências na



capacidade de acetilação da enzima N-acetiltransferase 2 (devido a alterações genéticas) se apresentam como um fator de risco genético para o aparecimento da lesão hepática.

A isoniazida é geralmente administrada de forma oral sendo rapidamente absorvida através do trato gastrointestinal. Primeiramente, através do sistema porta-hepático, o fármaco passa pelo fígado antes que possa atingir a circulação e antes que seja distribuído ao seu local de ação. Sendo assim, como a isoniazida é metabolizada por enzimas hepáticas, parte dela é inativada inicialmente, e sua biodisponibilidade diminui (efeito de primeira passagem). Cerca de 75% a 95% de uma dose de isoniazida são excretados na urina em 24 horas, principalmente na forma de metabólitos, sendo os principais produtos de excreção aqueles resultantes de acetilação enzimática (acetil-isoniazida) e hidrólise enzimática (ácido isonicotínico) (Hardman *et al.*, 1996).

A principal enzima responsável pela inativação e metabolização da INH é a N-acetiltransferase 2 (NAT2), codificada pelo gene *NAT2*. NAT2 é uma enzima de conjugação de Fase II que catalisa a transferência do grupo acetil, a partir do cofator Acetil-coenzima A (Acetil-CoA), ao terminal nitrogenado da droga (Figura 8). Após acetilação, a acetil-isoniazida é hidrolizada por uma amidase à acetil-hidrazina. Esta, por sua vez, é oxidada em intermediários hepatotóxicos pela mono-oxigenase CYP450 2E1. Paralelamente, a acetil-hidrazina é novamente acetilada por NAT2 formando diacetil-hidrazina, composto não tóxico. A acetil-hidrazina é considerada uma hepatotoxina, pois sua ativação metabólica por mono-oxigenases hepáticas (CYP2E1) gera metabólitos acilantes e alquilantes altamente reativos. A ligação covalente desses agentes (considerados reativos eletrofílicos ou radicais livres) a macromoléculas no fígado é capaz de causar lesão hepática (Figura 8) (Mitchell, *et al.*, 1976; Nelson, *et al.*, 1976; Woodward & Timbrell, 1984; Kita *et al.*, 2001; Huang *et al.*, 2002; Huang *et al.*, 2003; Huang, 2007.).



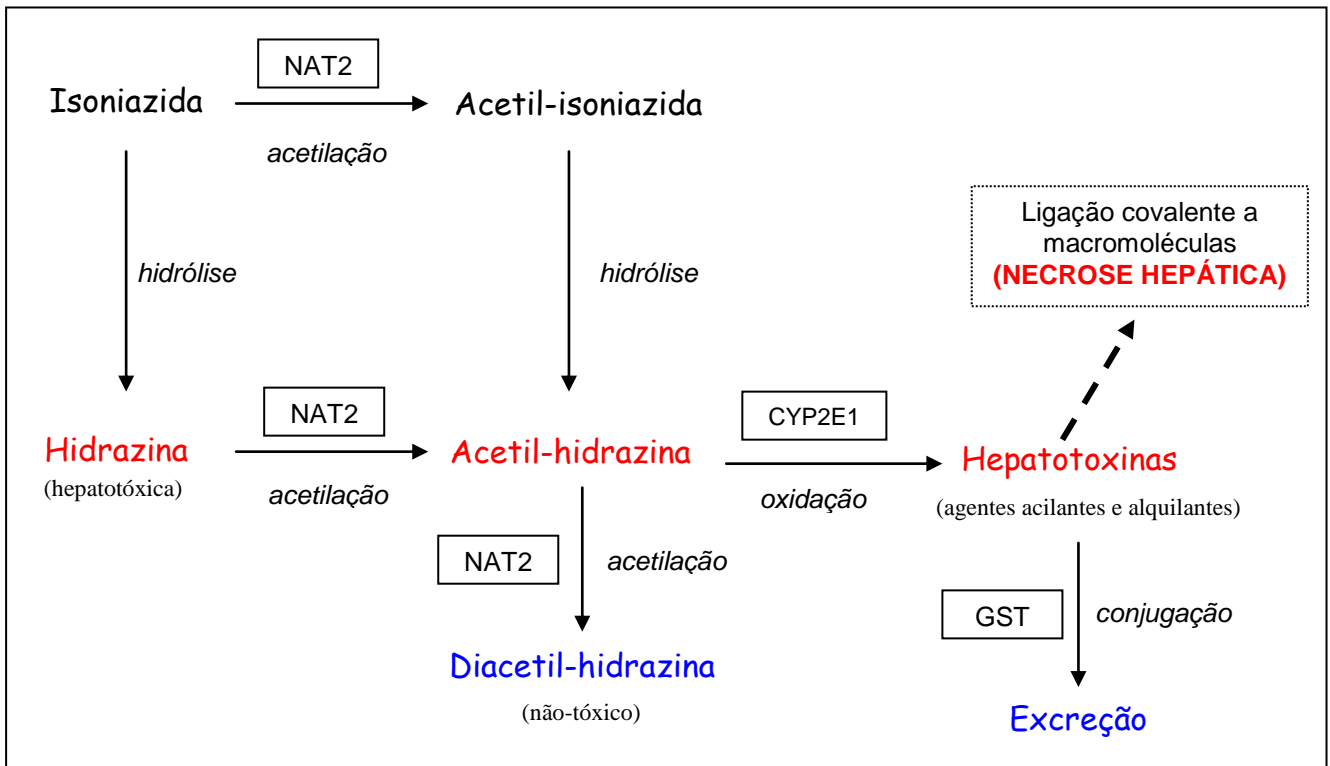


Figura 8: Metabolismo da isoniazida no fígado. Adaptado a partir de Mitchell, *et al.*, 1976 e Huang, 2007.

Mutações pontuais na região codificante de *NAT2* são capazes de alterar a atividade de acetilação da enzima e essas alterações podem gerar três fenótipos diferentes: acetiladores lentos, intermediários e rápidos. Os indivíduos que apresentam uma capacidade de acetilação lenta acumulam mais acetil-hidrazina, uma vez que a isoniazida compete com esse metabólito tóxico pela acetilação. A reação da acetil-hidrazina (em altos níveis) com *CYP2E1* leva a ligação covalente de metabólitos secundários altamente reativos com proteínas intracelulares. A partir daí, ocorrem alterações intracelulares que resultam na perda de gradientes iônicos, queda dos níveis de ATP e ruptura da actina com consequente lise celular (Figura 6). Em 1986, Sarma e colaboradores propuseram que a interação medicamentosa de INH com rifampicina (RIF) pode também levar à lesão hepática. Segundo os autores, a RIF aumenta a conversão direta da isoniazida a ácido isonicotínico e hidrazina (hepatotóxica, carcinogênica e neurotóxica) em acetiladores lentos, provocando o acúmulo desse metabólito hepatotóxico (Sarma *et al.*, 1986).

Outra classe de enzimas cuja participação na biotransformação da isoniazida que vem sendo estudada é a glutatona S-transferase (GST) (Roy, *et al.*, 2001; Huang, 2007) que realiza uma reação de conjugação da glutatona (GSH) a compostos eletrofílicos, levando à formação de produtos menos reativos e prontos para excreção pelo organismo (Figura 8).



### 1.5) A ciência da Farmacogenética

A variabilidade interindividual no metabolismo de xenobióticos e na resposta a uma determinada droga é extensa. Fatores que influenciam essa variação incluem: herança genética, sexo, idade, co-administração de medicamentos, fisiologia do indivíduo, presença de outra fisiopatologia e alguns fatores ambientais (consumo de álcool, cigarro, hábitos alimentares, etc).

Para produzir seus efeitos terapêuticos, um fármaco precisa estar presente em concentrações adequadas em seus locais de ação. Embora as concentrações obtidas sejam dependentes da dose administrada, elas também irão depender da magnitude e velocidade de sua absorção, distribuição, ligação ou localização nos tecidos, biotransformação e excreção. É a farmacocinética a ciência que estuda o caminho e distribuição da droga e seus metabólitos nos diferentes tecidos, englobando os mecanismos de absorção, transporte, metabolização e eliminação. O estudo dos efeitos bioquímicos e fisiológicos dos fármacos e seus mecanismos de ação é denominado farmacodinâmica. Proteínas envolvidas na mediação dos efeitos da droga são definidas como moléculas-alvo, e incluem não somente receptores diretos da mesma e, sim, qualquer proteína envolvida no seu mecanismo de ação (proteínas transdutoras de sinal, por exemplo). (Hardman *et al.*, 1996).

Após sua administração, a droga deve ser absorvida pelo organismo e distribuída por todo o corpo para que possa atingir seu sítio de ação. Isso requer a participação em conjunto de várias proteínas, incluindo enzimas metabolizadoras, proteínas transportadoras, receptores e outros. A droga, por sua vez, pode ser administrada na sua forma ativa ou inativa (pró-droga). Ambas as formas sofrem o processo de biotransformação para serem eliminadas do organismo e/ou liberarem seus componentes ativos, processo onde outras proteínas enzimáticas estão envolvidas (Prows & Prows, 2004). A variabilidade genética já é conhecida nos contextos de absorção e metabolismo de fármacos, além de suas interações com receptores farmacológicos (moléculas-alvo), levando então ao aparecimento de diferentes fenótipos de absorção (lenta ou rápida), metabolização (ineficiente, eficiente ou ultrarrápida) e ainda de interação droga-receptor (fraca ou ineficiente) (Ingelman-Sundbug, 2001). Uma das consequências dessas variações é o surgimento de reações adversas ou falência terapêutica (Figura 9).

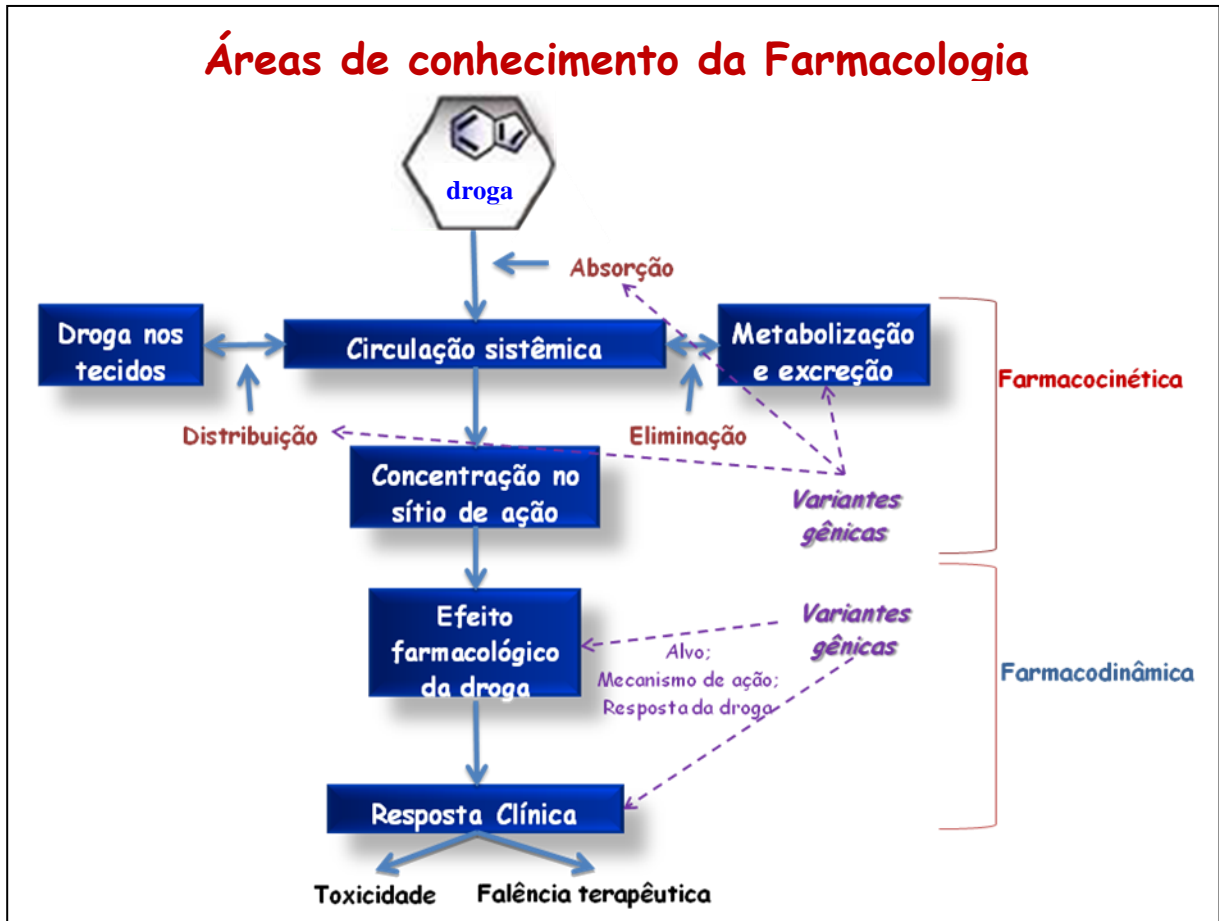


Figura 9: Componentes importantes da farmacologia. Variantes gênicas envolvidas com a farmacodinâmica ou farmacocinética da droga são alvos dos estudos farmacogenéticos.

A farmacogenética é o estudo das variações genéticas que causam respostas variáveis aos medicamentos incluindo diferenças na eficácia, interações droga-droga e o risco relativo de reações adversas. Isso envolve, então, estudos de mutações e polimorfismos que podem afetar a expressão ou atividade de receptores farmacológicos, proteínas transportadoras, sinalizadoras e de enzimas metabolizadoras (Ingelman-Sundberg, 2001; Roses, 2002, Prows & Prows, 2004).

Um dos principais objetos de estudo da farmacogenética são as enzimas envolvidas na biotransformação de fármacos. Variações genéticas nos genes que codificam essas enzimas podem resultar em alterações na atividade biotransformadora. Isso favoreceria o aparecimento de reações adversas por toxicidade devido a uma baixa metabolização e o consequente acúmulo do princípio ativo ou de metabólitos intermediários tóxicos (metabolizadores lentos). A presença de variantes que não afetam a atividade enzimática (metabolizadores rápidos), ou com múltiplas cópias (metabolizadores ultrarrápidos), podem estar associadas a concentrações séricas subterapêuticas da droga (Ingelman-Sundberg, 2001). Na maioria dos casos, a



presença de heterozigose origina metabolizadores intermediários (Kalow, 1997) (Figuras 10 e 11).

Cerca de 99,9% do genoma humano é idêntico em todos os indivíduos. Entretanto, ao longo da sequência deste genoma existem locais específicos onde ocorrem variações no DNA. Uma dessas alterações é conhecida como SNP (*Single Nucleotide Polymorphism*) e envolve a troca de uma base nucleotídica por outra qualquer (mutação de ponto).

Os dois tipos principais de mutação de ponto no DNA são as substituições de bases e adições ou deleções de bases (*indel*). As substituições de base podem ser de dois tipos: transições, quando ocorre substituição de uma base por outra da mesma categoria química – purina por purina (A/G) ou pirimidina por pirimidina (T/C); e transversões, quando ocorre substituição de uma base por outra de categoria química diferente – purina por pirimidina ou pirimidina por purina. As mutações de ponto podem ainda ser classificadas como: mutações sinônimas ou silenciosas (mutações que alteram um determinado códon para outro códon que codifica o mesmo aminoácido); mutações não sinônimas ou de sentido trocado (o códon para um determinado aminoácido é trocado por outro que determina aminoácido diferente); mutações sem sentido (o códon que codifica para um determinado aminoácido é trocado para um códon de terminação da tradução) e mutações de troca de matriz de leitura ou *frame-shift* (mutações *indel* mudam a matriz de leitura para o restante do processo de tradução) (Guttmacher & Collins, 2002).

Os SNPs são comuns e ocorrem a uma taxa de aproximadamente um, a cada trezentos a mil bases. Um conjunto de SNPs em um dado indivíduo é chamado de perfil de SNPs. Esses perfis estão sendo citados como alternativa para investigar as bases genéticas de várias doenças e ajudar a prever que tipo de resposta um determinado paciente pode apresentar em relação a uma determinada droga (variantes associadas com eficácia terapêutica e efeitos adversos) (Roses, 2002; Prows & Prows, 2004).

A identificação de marcadores genéticos que possam prever se o indivíduo responde mal, ou não, a uma determinada droga auxiliaria na indicação correta do medicamento e do seu esquema posológico, maximizando a eficácia e prevenindo ou reduzindo as possíveis reações adversas tóxicas (Figura 12).

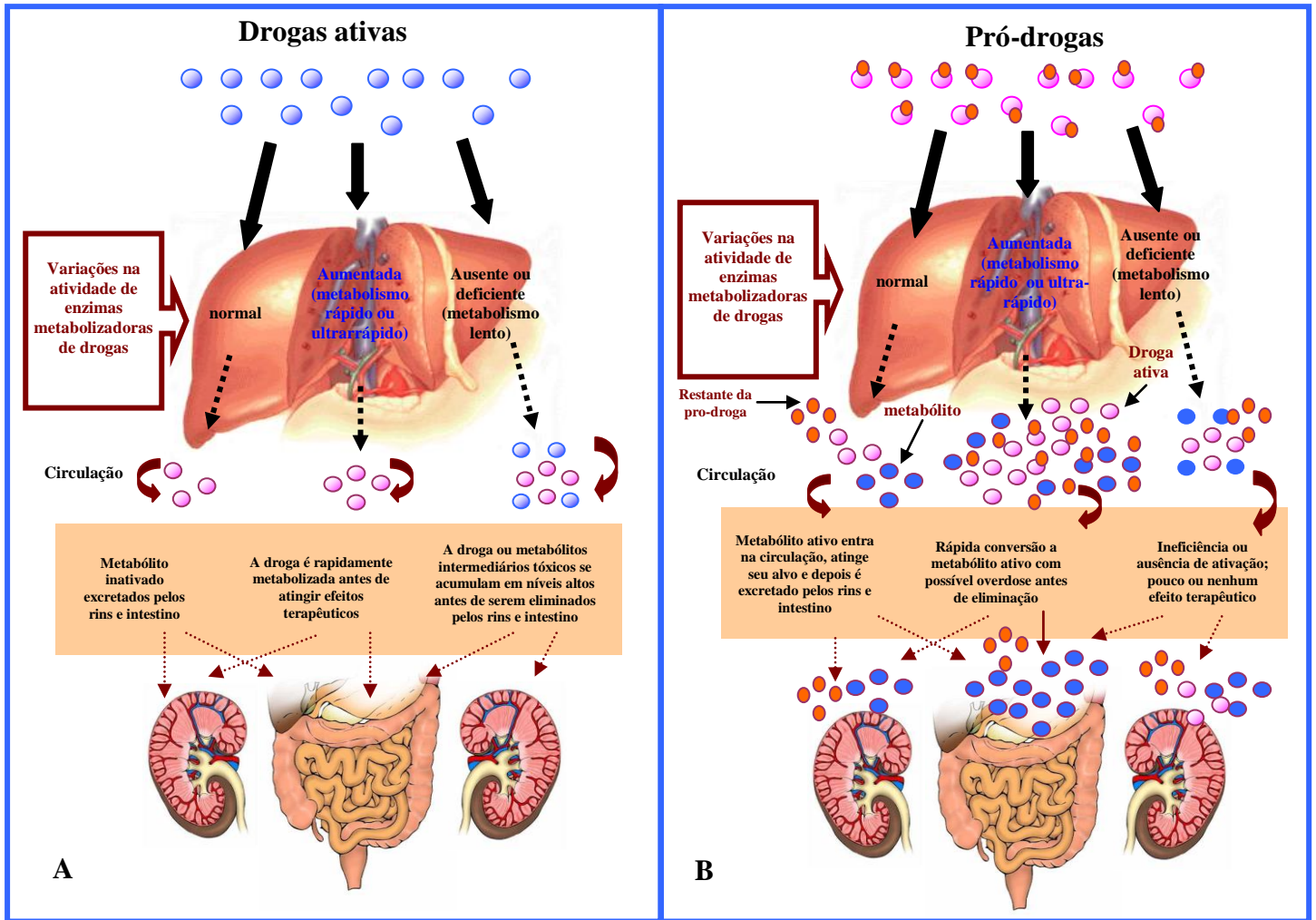


Figura 10: Esquema mostrando as três possibilidades resultantes do metabolismo de fármacos administrados como droga ativa (A) ou pró-droga (B) (Prows & Prows, 2004). Essas diferenças na biotransformação das drogas são definidas pelo genótipo do paciente. Os desfechos encontrados no quadro B são referentes ao metabolismo de pró-drogas que são ativadas por enzimas do indivíduo; a isoniazida, apesar de ser considerada uma pró-droga, se encaixa no esquema do quadro A, uma vez que sua ativação acontece dentro da micobactéria.

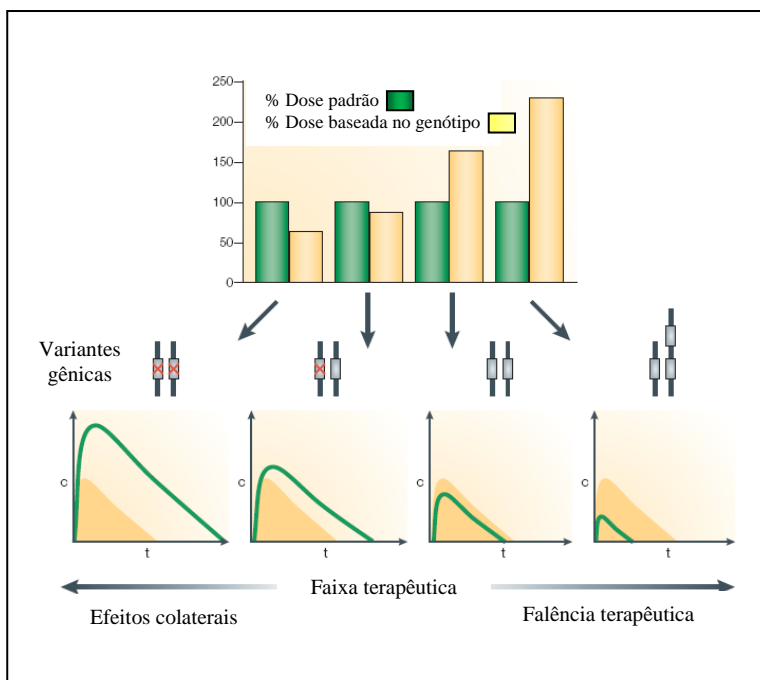


Figura 11: Esquema mostrando o benefício do ajuste de dose de uma droga, baseado no genótipo do indivíduo. Indivíduos com variantes gênicas que codificam enzimas metabolizadoras de drogas com baixa atividade têm maior propensão a desenvolver efeitos colaterais. Sendo assim, a dosagem do fármaco para esses indivíduos deve ser diminuída. Por outro lado, aqueles indivíduos que possuem genótipos de metabolizadores rápidos, ou ultrarrápidos, correm o risco de apresentar falha terapêutica devendo, então, receber doses aumentadas do fármaco. Adaptado de Kirchheiner *et al.*, 2005. c = concentração da droga; t = tempo.

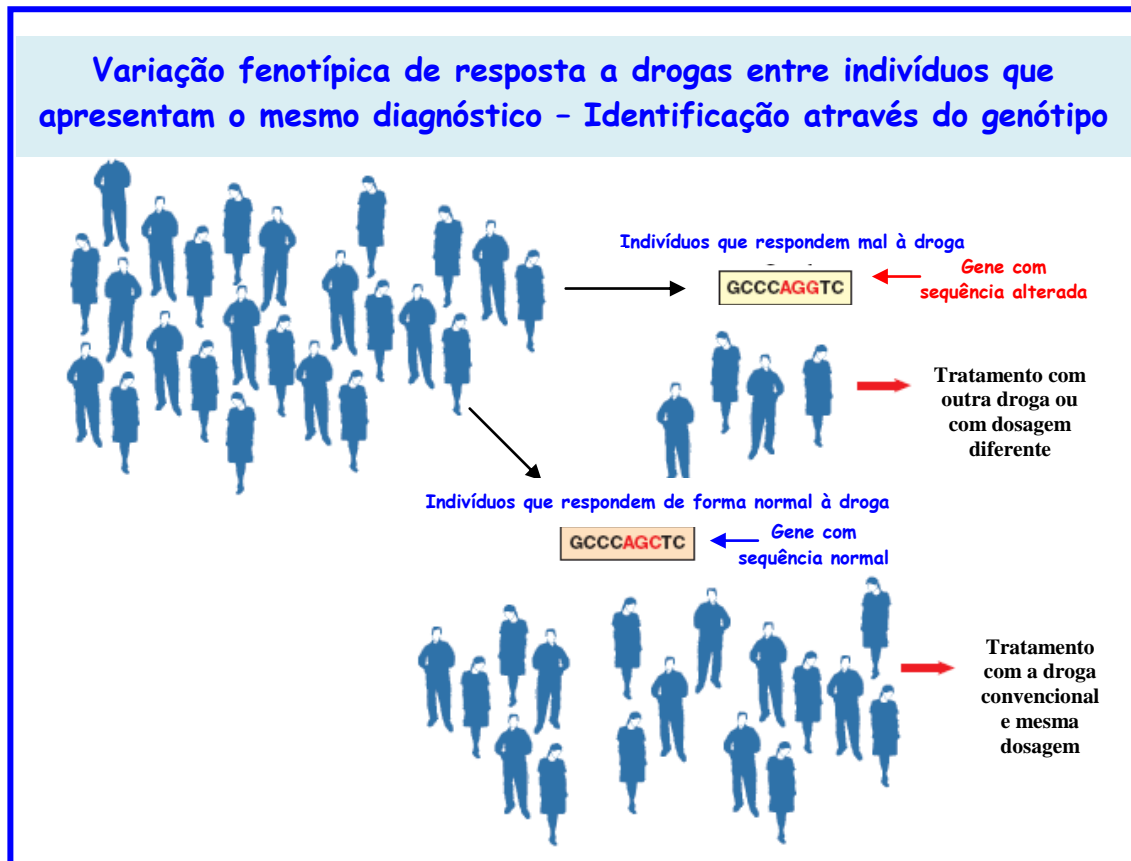


Figura 12: Pacientes que apresentam o mesmo esquema terapêutico podem responder de forma diferenciada, devido à presença de alelos diferentes. A prática da farmacogenética envolve a genotipagem de pacientes e o tratamento daqueles que respondem mal à quimioterapia de primeira escolha com outra droga, ou com uma posologia diferente (Prows & Prows, 2004).

## 1.6) Genética molecular da N-acetiltransferase 2

Variações na atividade da enzima N-acetiltransferase 2 foram descobertas há mais de 40 anos, através da observação de diferenças interindividuais no metabolismo da isoniazida, em pacientes com tuberculose, bem como da presença ou não de toxicidade induzida por essa droga (Evans, 1989). Alguns estudos de genética de família puderam então comprovar que a variabilidade genética de *NAT2* estava diretamente relacionada ao aparecimento de diferentes fenótipos de acetilação (Knight *et al.*, 1959).

O estudo molecular das N-acetiltransferases em humanos revelou a presença de três *loci*, dois que codificam para duas enzimas similares (*NAT1* e *NAT2*) e um terceiro que constitui um pseudogene (*pNAT*) (Figura 13). Os três *loci* estão localizados no cromossomo 8 entre 170-360Kb da região 8p22, uma região instável e frequentemente deletada em células tumorais, sugerindo a presença de genes supressores de tumor nesse local (Sim *et al.*, 2000). O *pNAT* é um pseudogene que contém um códon de parada prematuro, e não é transcrito.



*NAT1* e *NAT2* consistem em dois genes de 873 pb, que não apresentam íntrons, e codificam proteínas de 34 KDa. Enquanto a proteína *NAT1* é derivada de um mRNA sintetizado a partir de um único éxon, o mRNA de *NAT2* é constituído do éxon codificador da proteína e um segundo éxon, não codificador, de 100 pb localizado a aproximadamente 8Kb antes do sítio de iniciação da tradução. *NAT1* e *NAT2* apresentam 81% de homologia e seus respectivos genes apresentam 87% de homologia na sequência nucleotídica. Ambas as enzimas são capazes de realizar a N-acetilação, O-acetilação e N-O-transferência em diferentes xenobióticos e carcinógenos; entretanto, diferem na distribuição tecidual no organismo, na especificidade a substratos e em níveis de expressão durante o desenvolvimento embrionário (Tabela 2). (Hein, *et al.*, 2000; Sim *et al.*, 2000; Hein, 2002; Pompeo *et al.*, 2002).

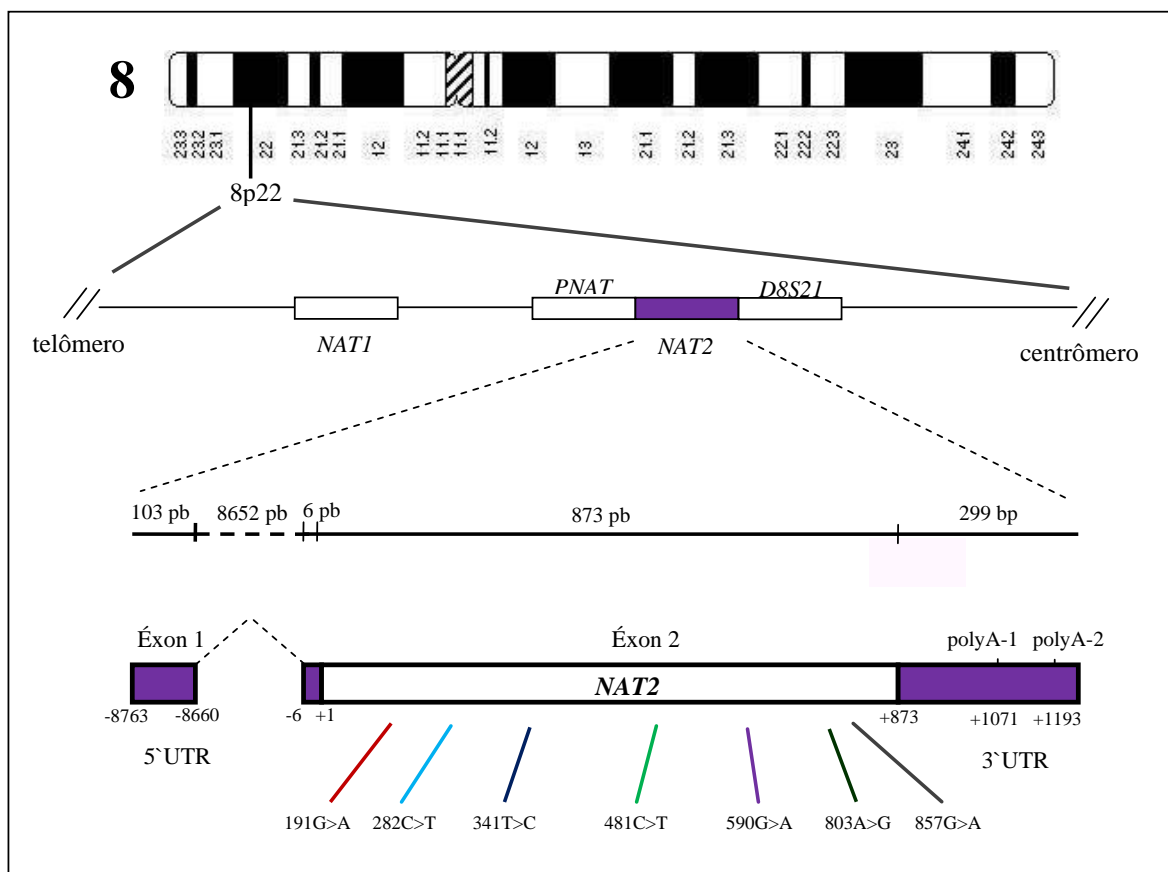


Figura 13: Localização cromossômica dos loci das N-acetiltransferases humanas. Esquema mostrando os SNPs mais frequentes encontrados no gene *NAT2* humano. *D8S21* representa um marcador polimórfico encontrado dentro no locus de *NAT2* (Adaptado a partir de Sim *et al.*, 2000 e Boukouvala & Sim, 2005).

Tanto *NAT1* como *NAT2* são genes polimórficos com mais de 20 alelos descritos para cada locus. Uma vez que o polimorfismo genético influencia na velocidade de acetilação de *NAT1* e *NAT2*, os indivíduos são então classificados em acetiladores rápidos (e/ou intermediários) e lentos. Apesar de alguns estudos mostrarem que o fenótipo de acetilação lenta obedece a uma herança autossômica recessiva (Knight *et al.*, 1959; Jorge-Nebert *et al.*,



2002; Moussa *et al.*, 2002; Agúndez *et al.*, 2002; Dandara *et al.*, 2003; Mittal *et al.*, 2004), outros estudos mostram a presença de três genótipos codominantes: homozigotos rápidos, heterozigotos intermediários e homozigotos lentos (Sanahara *et al.*, 1961; Mashimo *et al.*, 1992; Doll *et al.*, 1995; Deitz *et al.*, 2000; Ohno *et al.*, 2000; Hein *et al.*, 2000; Kita *et al.*, 2001; Pompeo *et al.*, 2002). As variações entre esses estudos possivelmente se devem à utilização de diferentes substratos (INH, dapsona, cafeína, procainamida, ácido aminossalicílico, sulfapiridina, etc) nos ensaios enzimáticos de fenotipagem, uma vez que algumas dessas substâncias não são capazes de discriminar os acetiladores rápidos dos intermediários (Mashimo *et al.*, 1992; Yokogawa *et al.*, 2001). De fato, alguns trabalhos onde foram realizados experimentos de genotipagem e avaliação do fenótipo concomitantemente mostraram que o comportamento dos heterozigotos (ditos por alguns autores heterozigotos rápidos) difere de ambos os homozigotos, apresentando uma atividade enzimática intermediária (Deguchi, *et al.*, 1990; Hickman & Sim, 1991; Kita *et al.*, 2001; Jorge-Nebert *et al.*, 2002). Adicionalmente, quando os três genótipos foram avaliados quanto aos níveis séricos atingidos pela INH horas depois da sua administração, novamente o heterozigoto se manteve numa faixa intermediária em relação aos homozigotos rápidos e lentos (Donald, *et al.*, 2004). Entretanto, essas discrepâncias ainda não foram resolvidas de forma satisfatória.

Tabela 2: Características moleculares das N-acetiltransferases humanas e a relação entre o fenótipo e genótipo apresentados por essas enzimas (Hein, *et al.*, 2000; Sim *et al.*, 2000; Hein, 2002; Pompeo *et al.*, 2002).

CARACTERÍSTICA	N-ACETILTRANSFERASE 1	N-ACETILTRANSFERASE 2
Localização cromossômica	8p22 (170-360 Kb) Região instável em tumores (presença de gene supressor de tumor)	
Gene	NAT1 (873 pb), sem íntrons; NAT2 (873 pb), sem íntrons (mRNA composto pelo éxon codificador da proteína e éxon de 100pb não codificante localizado a 8 Kb a cima do códon de iniciação)	
Proteína	NAT1 (290 a.a. - 34 KDa)	NAT2 (290 a.a. - 34 KDa)
Atividade enzimática	N-acetilação (desintoxicação de xenobióticos) O-acetilação (ativação de carcinógenos) N-O-acetiltransferência(ativação de carcinógenos)	
Homologia	gene= 87% proteína = 81%	
Locais de expressão	Diferentes tecidos Fases iniciais do desenvolvimento embrionário e fase adulta	Epitélio intestinal e fígado (mRNA encontrado em diferentes tecidos)
Substratos	Aminas heterocíclicas, ácido p-aminossalicílico, ácido p-aminobenzoico, p-aminobenzoil glutamato (p-AB Glu) → substrato endógeno	Isoniazida, sulfonamidas (sulfametazina), dapsona, aminas heterocíclicas (PhIP, Glu-P1) derivadas de alimentos





Tabela 2: Características moleculares das N-acetiltransferases humanas e a relação entre o fenótipo e genótipo apresentados por essas enzimas (Hein, *et al.*, 2000; Sim *et al.*, 2000; Hein, 2002; Pompeo *et al.*, 2002). (Continuação)

CARACTERÍSTICA	N-ACETILTRANSFERASE 1	N-ACETILTRANSFERASE 2
Polimorfismo genético	NAT1 = mais de 20 alelos / NAT2 = 53 alelos Frequência dos alelos varia de acordo com grupos étnicos Alelo = combinação de mutações pontuais no promotor e/ou no CDS / SNPs, encontrados nos diferentes alelos, resultam em redução de afinidade pelo substrato, atividade enzimática, níveis da proteína e/ou estabilidade da mesma.	
Genótipo/fenótipo	TRIMODAL: acetiladores lentos, intermediários e rápidos. Relação com predisposição ao câncer e/ou eficácia ou efeitos colaterais no tratamento com antibióticos	

Análises funcionais e estudos de mutação sitiodirigida permitiram a identificação de um resíduo de cisteína no sítio ativo (na posição 68 em humanos – Cys68) como a única cisteína conservada em todas as NATs cujas sequências já estavam disponíveis até o momento. Esses estudos demonstraram que Cys68 desempenha um papel crucial no sítio catalítico da enzima NAT2 humana (Watanabe *et al.*, 1992, Dupret & Grant, 1992). Posteriormente, estudos de modelagem molecular por homologia da NAT2 humana, baseados na estrutura protéica de NAT *Salmonella typhimurium*, revelaram a presença de uma tríade catalítica similar à encontrada em cisteína-proteases (Cys68-His107-Asp122) (Sinclair *et al.*, 2000, Rodrigues-Lima & Dupret, 2002).

Os estudos baseados na estrutura cristalográfica de NATs de diferentes bactérias foram de grande importância para o melhor entendimento das características estruturais dessas proteínas, incluindo a localização de SNPs ao longo da estrutura tridimensional e experimentos de *docking* molecular com diversos substratos (Sinclair *et al.*, 2000; Sandy *et al.*, 2002; Holton *et al.*, 2005; Westwood *et al.*, 2005; Westwood *et al.*, 2006). Entretanto, a similaridade entre as sequências protéicas de NATs de mamíferos e NATs bacterianas é muito baixa (30%), e as diferenças estruturais entre essas proteínas representam um fator importante e limitante para o estudo de modelagem molecular das N-acetiltransferases de seres humanos (Walraven *et al.*, 2007).

Recentemente, as estruturas cristalográficas de NAT1 e NAT2 humanas foram resolvidas, corroborando vários achados prévios de estudos de modelagem molecular e fornecendo evidências da existência de novas características estruturais na porção C-terminal das NATs humanas ausentes em bactérias (Wu *et al.*, 2007). Adicionalmente, foi verificado que o arranjo tridimensional de NAT1 e NAT2 humanas é muito similar aos seus ortólogos bacterianos, com três domínios principais: um domínio N-terminal de oitenta e três resíduos



contendo cinco  $\alpha$ -hélices e uma folha- $\beta$  pequena; um segundo domínio com nove folhas- $\beta$  e duas pequenas hélices formando um barril- $\beta$ ; e um terceiro domínio  $\alpha/\beta$  com quatro folhas- $\beta$  e uma  $\alpha$ -hélice (Figura 14). Os dois primeiros domínios, envolvidos na atividade catalítica e ligação do Acetil-CoA, estão conectados ao terceiro domínio através de um interdomínio  $\alpha$ -helicoidal. O terceiro domínio está relacionado com a afinidade ao substrato (Westwood *et al.*, 2006; Wu *et al.*, 2007).

A N-acetilação envolve uma reação clássica do tipo “ping-pong”, em duas etapas, e utiliza o cofator Acetil-CoA como doador do grupamento acetil. O resíduo sulfidril da cisteína no sítio catalítico de NAT recebe o grupamento acetil, num primeiro momento da reação, resultando na formação da estrutura intermediária acetilcisteinil. Posteriormente, o grupamento acetil é transferido para o terminal amino da molécula aceptora. (Dupret & Grant, 1992, Wang *et al.*, 2004).

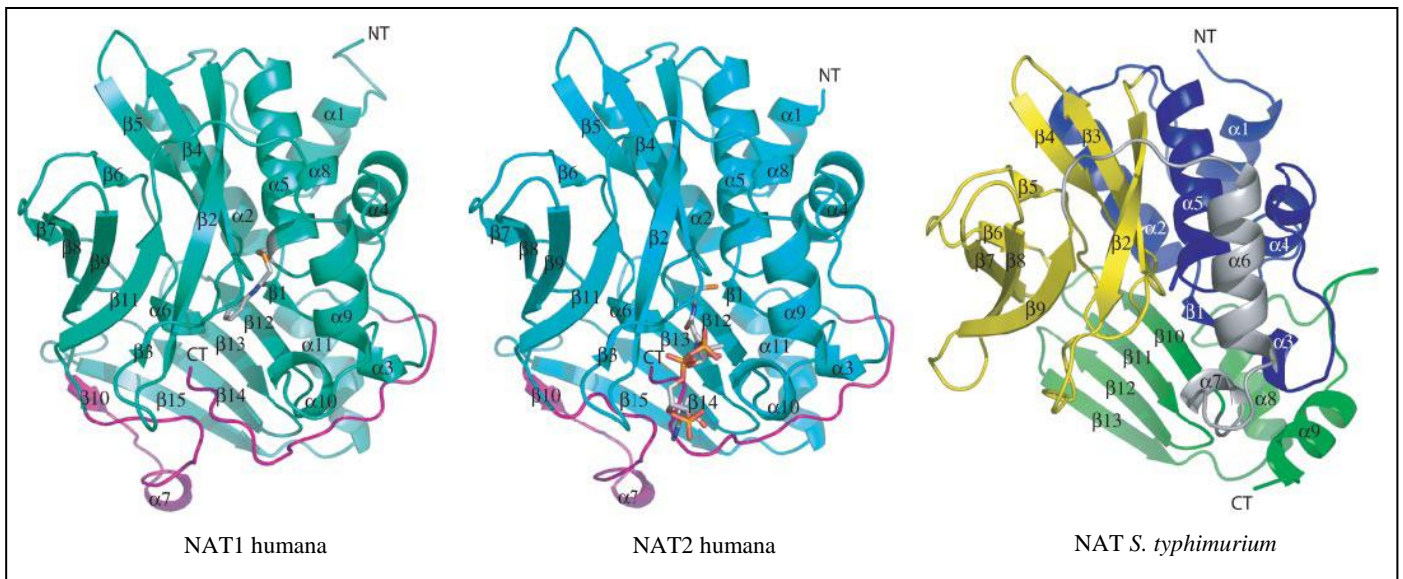


Figura 14: Estrutura geral das N-acetiltransferases 1 e 2 humanas e de *S. typhimurium*. A NAT1 humana está ligada ao substrato acetamina e a NAT2 humana está ligada a Acetil-CoA. Os terminais amino e carboxil estão indicados como NT e CT, respectivamente. A inserção de uma cauda na região carboxi-terminal das NATs humanas, inexistente na NAT bacteriana, está destacada em roxo. Os três domínios estão representados por cores na NAT de *S. typhimurium*: domínio I = azul; domínio 2 = amarelo; domínio III = verde. O interdomínio está destacado em cinza. Figura retirada de Wu *et al.*, 2007.

Até o presente momento, já foram identificados cinquenta e três alelos diferentes para o gene *NAT2* ([www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html)), os quais consistem na combinação de até quatro mutações pontuais na região codificante (Tabela 3). Existem trinta e dois SNPs identificados na região codificante de *NAT2*: vinte e cinco cadastrados no site oficial de nomenclatura e sete no banco de dados SNPentrez e/ou em artigos publicados. Estudos funcionais de expressão gênica, em células de levedura e de mamífero, mostraram claramente a existência de diferentes mecanismos que podem resultar no fenótipo de



acetilação lenta de NAT2 (Fretland *et al.*, 2001; Walraven *et al.*, 2008). Nesses estudos, foi observado que as mutações no gene *NAT2* podem interferir na quantidade de proteína sintetizada (190C>T, 191G>A, 341T>C, 364G>A, 411A>T, 434A>C, 499G>A, 590G>A e 857G>A), na atividade enzimática (190C>T, 191G>A, 341T>C, 364G>A, 411A>T, 434A>C, 499G>A, 590G>A, 845A>C e 857G>A) e estabilidade de NAT2 (191G>A, 590G>A, 845A>C e 857G>A), conferindo a todos um fenótipo de acetilação lenta. Entretanto, outras mutações (111T>C, 282C>T, 481C>T, 759C>T e 803A>G) não exerceram nenhum efeito na proteína, resultando em um fenótipo normal, semelhante a proteína codificada pelo alelo tipo-selvagem (Fretland *et al.*, 2001; Walraven *et al.*, 2008).

Tabela 3: Alelos de *NAT2* já descritos na literatura. As mutações se referem à posição dentro da região codificante de *NAT2*, considerando o nucleotídeo de número um (1) a adenina (A) do códon de iniciação da tradução (ATG). A mutação que caracteriza cada grupo está destacada em vermelho.

ALELOS	SNPs	ALELOS	SNPs	ALELOS	SNPs
<i>NAT2*4</i>	tipo-selvagem	<i>NAT2*6E</i>	481C>T 590G>A	<i>NAT2*12F</i>	622T>C 803A>G
<i>NAT2*5A</i>	341T>C 481C>T	<i>NAT2*6F</i>	590G>A 803A>G	<i>NAT2*12G</i>	609G>T 803A>G
<i>NAT2*5B</i>	341T>C 481C>T 803A>G	<i>NAT2*6G</i>	282C>T 518A>G 590G>A	<i>NAT2*12H</i>	403C>G 803A>G
<i>NAT2*5C</i>	341T>C 803A>G	<i>NAT2*6H</i>	282C>T 590G>A 766A>G	<i>NAT2*13A</i>	282C>T
<i>NAT2*5D</i>	341T>C	<i>NAT2*6I</i>	282C>T 590G>A 838G>A 857G>A	<i>NAT2*13B</i>	282C>T 578C>T
<i>NAT2*5E</i>	341T>C 590G>A	<i>NAT2*6J</i>	282C>T 590G>A 857G>A	<i>NAT2*14A</i>	191G>A
<i>NAT2*5F</i>	341T>C 481C>T 759C>T 803A>G	<i>NAT2*6K</i>	282C>T 590G>A 638C>T	<i>NAT2*14B</i>	191G>A 282C>T
<i>NAT2*5G</i>	282C>T 341T>C 481C>T 803A>G	<i>NAT2*6L</i>	282C>T 345C>T 590G>A	<i>NAT2*14C</i>	191G>A 341T>C 481C>T 803A>G
<i>NAT2*5H</i>	341T>C 481C>T 803A>G 859T>C	<i>NAT2*7A</i>	857G>A	<i>NAT2*14D</i>	191G>A 282C>T 590G>A
<i>NAT2*5I</i>	341T>C 411A>T 481C>T 803A>G	<i>NAT2*7B</i>	282C>T 857G>A	<i>NAT2*14E</i>	191G>A 803A>G
<i>NAT2*5J</i>	282C>T 341T>C 590G>A	<i>NAT2*10</i>	499G>A	<i>NAT2*14F</i>	191G>A 341T>C 803A>G
<i>NAT2*5K</i>	282C>T 341T>C	<i>NAT2*11A</i>	481C>T	<i>NAT2*14G</i>	191G>A 282C>T 803A>G



Tabela 3: Alelos de *NAT2* já descritos na literatura. As mutações se referem à posição dentro da região codificante de *NAT2*, considerando o nucleotídeo de número um (1) a adenina (A) do códon de iniciação da tradução (ATG). A mutação que caracteriza cada grupo está destacada em vermelho. (Continuação)

ALELOS	SNPs	ALELOS	SNPs	ALELOS	SNPs
<i>NAT2*5L</i>	70T>A 341T>C 481C>T 803A>G	<i>NAT2*11B</i>	481C>T 859Del	<i>NAT2*14H</i>	191G>A 282C>T 683C>T
<i>NAT2*5M</i>	341T>C 481C>T 803A>G 838G>A	<i>NAT2*12A</i>	803A>G	<i>NAT2*14I</i>	191G>A 481C>T 803A>G
<i>NAT2*6A</i>	282C>T 590G>A	<i>NAT2*12B</i>	282C>T 803A>G	<i>NAT2*17</i>	434A>C
<i>NAT2*6B</i>	590G>A	<i>NAT2*12C</i>	481C>T 803A>G	<i>NAT2*18</i>	845A>C
<i>NAT2*6C</i>	282C>T 590G>A 803A>G	<i>NAT2*12D</i>	364G>A 803A>G	<i>NAT2*19</i>	190C>T
<i>NAT2*6D</i>	111T>C 282C>T 590G>A	<i>NAT2*12E</i>	282C>T 578C>T 803A>G		

Sabe-se hoje que, para os genes que contém vários SNPs, o haplótipo - combinação destes SNPs no mesmo *locus* - é o que determina o fenótipo, e não o SNP individual. Sendo assim, vários estudos de expressão e atividade enzimática foram realizados com os alelos mais frequentes de *NAT2*, possibilitando então a caracterização dos mesmos quanto ao perfil de acetilação: os alelos dos grupos *NAT2\*5*, *\*6*, *\*7*, e *\*14* resultam em um fenótipo de acetilação lenta e os alelos *NAT2\*4*, do grupo *\*12* (com exceção do alelo *\*12D*) e *\*13A* conferem um fenótipo de acetilação rápida (Hein *et al.*, 2000; Hein, 2002).

Estudos realizados em populações distintas mostraram que a distribuição dos alelos que codificam para *NAT2* apresenta diferenças de acordo com a etnia da população estudada. Por exemplo, a mutação 191G>A, que caracteriza o *cluster \*14*, está presente nas populações americanas de origem africana, e em populações africanas nativas, mas nunca foi identificada em populações eurodescendentes. Os alelos *NAT2\*5A*, *NAT2\*5B*, *NAT2\*5C*, *NAT2\*6A* e *NAT2\*7B* representam 99% dos alelos mutantes em acetiladores lentos da população eurodescendente, onde esse fenótipo apresenta uma frequência acima dos 50%. Já nas populações asiáticas, o fenótipo de acetilação rápida predomina, sendo representado pelo alelo *NAT2\*4* (Tabela 4).



Tabela 4: Comparação interétnica da distribuição dos alelos de *NAT2* e a frequência do fenótipo de acetilação lenta encontrada nas diferentes populações. A soma dos alelos não representa necessariamente 100% porque na maioria dos estudos os alelos raros não foram considerados. (Tabela construída a partir das informações obtidas de Meyer & Zanger, 1997; Jorge-Nebert *et al.*, 2002; Moussa *et al.*, 2002; Dandara *et al.*, 2003).

População	Acetiladores lentos	<i>NAT2*4</i>	<i>NAT2*5</i>	<i>NAT2*6</i>	<i>NAT2*7</i>	<i>NAT2*14</i>
<b>Eurodescendente</b>	40-70%	-	-	-	-	-
✓ Americanos brancos	-	24%	43%	31%	2%	0
✓ Europeus brancos	-	26%	46%	26%	2%	0
✓ Espanhóis brancos	-	22%	44%	26%	1%	1%
✓ Franceses	59%	-	-	-	-	-
<b>Americanos-africanos</b>	-	35%	30%	23%	5%	8%
<b>Africanos nativos</b>	-	-	-	-	-	-
✓ Gabonese	-	27%	40%	22%	2%	9%
✓ Tanzânia	49%	30,8%	33,8%	20,5%	3,4%	11,5%
✓ Venda	38%	23,4%	38,5%	21,9%	5,2%	10,9%
✓ Zimbábwea	52%	28,8%	30,7%	20,9%	5,8%	13,8%
<b>Marroquinos</b>	61,8%	-	-	-	-	-
<b>Gregos</b>	66,3%	-	-	-	-	-
<b>Orientais</b>	10-30%	-	-	-	-	-
✓ Japoneses	11,4%	67%	1%	22%	10%	0
✓ Chineses	-	53%	5%	30%	12%	0
✓ Coreanos	-	68%	2%	18%	11%	1%
<b>Ameríndios</b>	-	-	-	-	-	-
✓ Ngawbe	27,6%	72%	2,4%	0	23,3%	-
✓ Embera	39%	61%	9,9%	3,7%	22,8%	-

Devido ao fato de *NAT1* e *NAT2* catalisarem a ativação metabólica de aminas aromáticas e heterocíclicas carcinogênicas assim como estarem envolvidas na biotransformação de xenobióticos, polimorfismos genéticos em *NAT1* e *NAT2* podem modificar o risco de câncer relacionado à exposição a esses carcinógenos, bem como resultar também na ocorrência de reações adversas a drogas e eficácia terapêutica. Na tentativa de se estabelecer uma associação entre o perfil de acetilação e o desenvolvimento de uma determinada doença, estudos de coorte ou do tipo caso-controle têm sido realizados através da



utilização de ferramentas de fenotipagem e genotipagem. No caso do gene *NAT2*, já é clara a correlação entre o fenótipo de acetilação lenta e o desenvolvimento de câncer de bexiga, bem como a correlação entre o fenótipo de acetilação rápida com o desenvolvimento de câncer de cólon (Hein *et al.*, 2000).

Durante muitos anos, a isoniazida vem sendo apontada como uma droga hepatotóxica e estudos de associação entre o fenótipo de acetilação e a suscetibilidade a essas reações adversas vêm sendo realizados. Dois estudos iniciais feitos na população oriental verificaram a associação do fenótipo de acetilação lenta com a hepatotoxicidade induzida pela INH (Ohno *et al.*, 2000; Huang *et al.*, 2002). Através da genotipagem pela técnica de PCR-RFLP esses autores comprovaram que, dentre os pacientes submetidos ao tratamento quimioterápico anti-TB, aqueles que continham alelos para acetilação lenta apresentavam risco aumentado de desenvolvimento de hepatotoxicidade pela isoniazida. Recentemente, novos estudos em diferentes populações observaram que, de fato, indivíduos portadores de um genótipo associado à acetilação lenta estariam predispostos a desenvolverem hepatite medicamentosa causada por drogas anti-TB (Hiratsuka *et al.*, 2002; Kinzig-Schippers *et al.*, 2005; Shimizu *et al.*, 2006; Cho *et al.*, 2007; Higuchi *et al.*, 2007). Em contrapartida, outros autores observaram associação entre acetiladores rápidos e suscetibilidade à ocorrência de efeitos colaterais durante o tratamento anti-TB (Mitchell *et al.*, 1975) ou nenhuma relação entre o perfil de acetilação e a hepatotoxicidade induzida por essas drogas (Singh *et al.*, 1995a; Singh *et al.*, 1995b; Roy *et al.*, 2006; Vuilleumier *et al.*, 2006), mostrando que esses resultados discrepantes precisam ser elucidados.

### 1.7) Genética molecular da mono-oxigenase CYP2E1

O citocromo P450 é uma hemoproteína e constitui o complexo enzimático mais importante das reações de biotransformação de Fase I. As isoenzimas que compõem o complexo CYP450 participam do sistema de oxidação da célula e estão localizadas na membrana do retículo endoplasmático liso, preferencialmente das células do fígado. Entretanto, essas mono-oxigenases já foram identificadas em diversos tecidos como no intestino, pâncreas, cérebro, pulmão, rim, medula óssea, pele, ovários e testículos (Wijnen *et al.*, 2007).

As proteínas CYP450 estão agrupadas em famílias e subfamílias de acordo com a similaridade entre as sequências de aminoácidos; membros da mesma família possuem  $\geq 40\%$



de identidade na sequência aminoacídica e membros da mesma subfamília compartilham  $\geq$  55% de identidade (Figura 15).

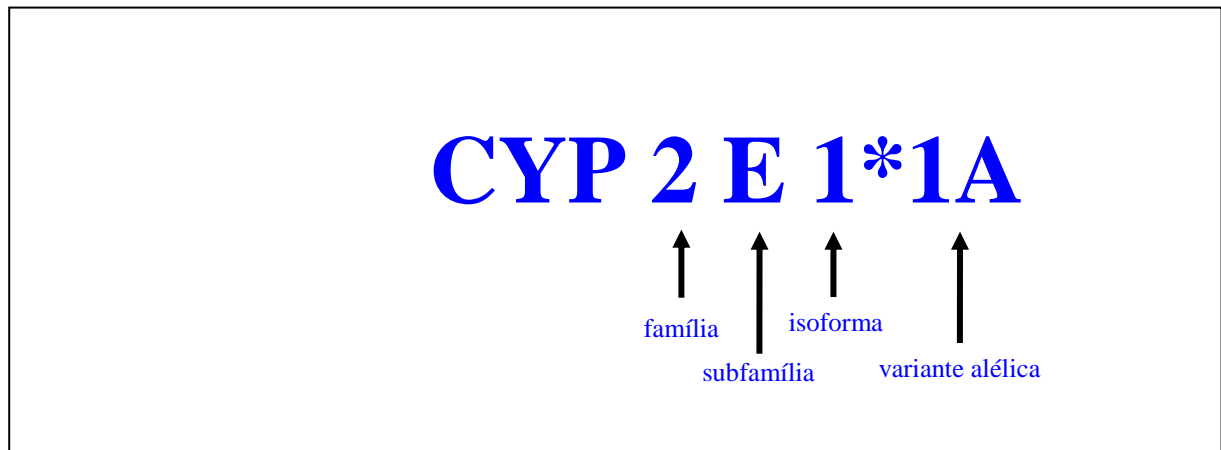


Figura 15: Nomenclatura das isoenzimas da super família CYP450.

As mono-oxigenases P450 atuam no metabolismo de vários substratos endógenos e na síntese de lipídios hidrofóbicos como o colesterol, hormônios esteróides, ácidos biliares e ácidos graxos. Além disto, algumas enzimas do complexo P450 metabolizam várias substâncias exógenas incluindo drogas, substâncias químicas ambientais e poluentes, bem como produtos derivados de plantas. O metabolismo de substâncias exógenas por CYP450 geralmente resulta na desintoxicação da substância; entretanto, as reações desencadeadas por essas enzimas podem levar a geração de metabólitos tóxicos que contribuem para o aumento do risco de desenvolvimento de cânceres e outros efeitos tóxicos (Nebert & Russel, 2002).

O sequenciamento do genoma humano permitiu a identificação de cento e quinze genes humanos de CYP450: cinquenta e sete genes ativos e cinquenta e oito pseudogenes (<http://drnelson.utmem.edu/cytochromeP450.html>). As isoenzimas CYP450 pertencentes às famílias 1-3 são responsáveis por 70-80% de todas as reações de Fase I dos fármacos utilizados atualmente, enquanto CYPs das outras famílias participam da metabolização de componentes endógenos (Wijnen *et al.*, 2007). A CYP2 constitui a maior família de isoenzimas e compreende um terço de todas as CYPs humanas (Figura 16). Os genes que codificam para essas enzimas são polimórficos e a frequência da distribuição das variantes alélicas nos diferentes grupos étnicos varia. Em geral, quatro fenótipos podem ser identificados: metabolizadores lentos, que não possuem uma boa atividade enzimática; metabolizadores intermediários, geralmente heterozigotos para um alelo deficiente; metabolizadores rápidos, que possuem dois alelos normais e, por último, os metabolizadores ultrarrápidos, que possuem várias cópias do gene (Ilgelman-Sundberg, 2004).

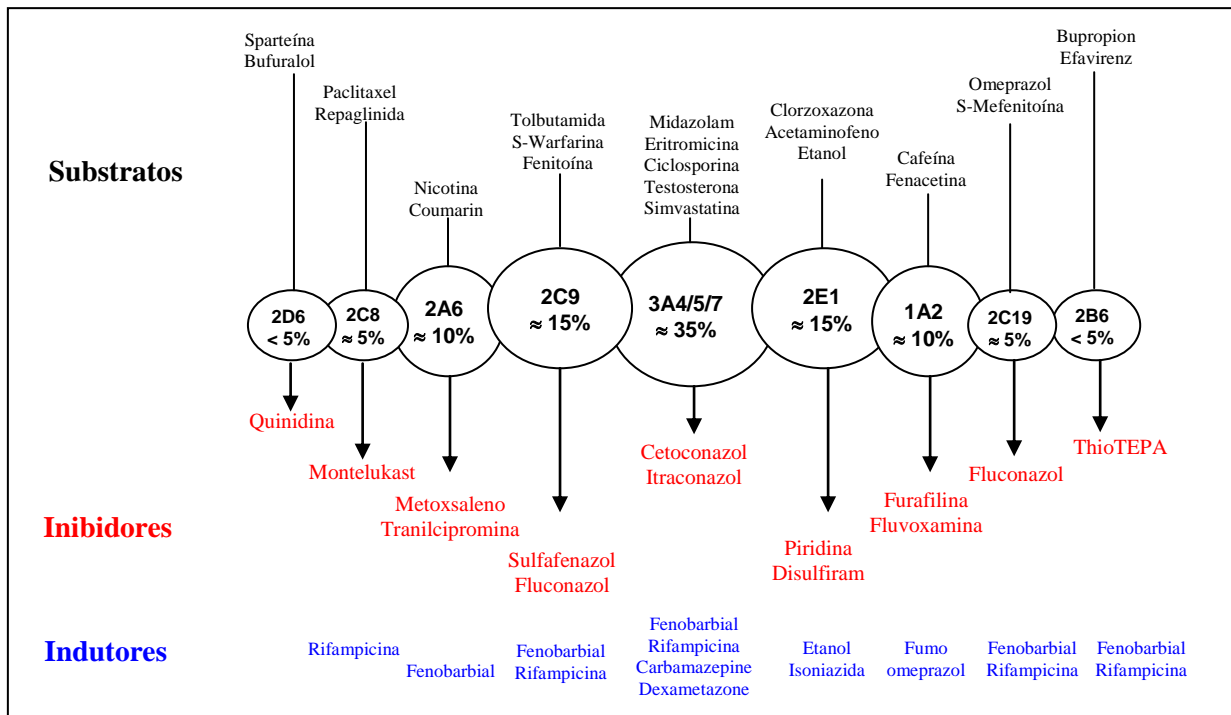


Figura 16: Abundância relativa das isoformas de CYP450 no fígado e alguns exemplos de substratos, inibidores e indutores (Pelkonen *et al.*, 2008).

A enzima CYP2E1, membro da superfamília do citocromo P450, é expressa principalmente no fígado, mas pode ser encontrada em outros órgãos como rins, trato gastrointestinal e cérebro. Suas funções abrangem desde a oxidação do etanol, gerando produtos reativos como acetaldeído e o radical 1-hidroxietil, até a metabolização de várias drogas e procarcinógenos. Além do etanol, CYP2E1 pode ser induzido por diversas drogas como, por exemplo, a isoniazida, e também por hidrocarbonetos, benzeno, clorofórmio e diversos solventes orgânicos (Caro & Cederbaum, 2004) (Figura 16).

A atividade de CYP2E1 é modulada pela presença de polimorfismos em vários sítios ao longo do gene. Dois polimorfismos se localizam na região promotora e são detectáveis através de digestão com as enzimas de restrição *Pst* I e *Rsa* I (Figura 17). Esses SNPs estão em completo desequilíbrio de ligação e podem estar relacionados com a expressão gênica de *CYP2E1*, uma vez que estão localizados no sítio de ligação do fator de transcrição HNF-1 (Watanabe *et al.*, 1994). Três genótipos de *CYP2E1* podem ser classificados pela técnica de PCR-RFLP utilizando a enzima *Rsa* I (c1/c1, c1/c2 e c2/c2), onde o alelo tipo-selvagem é denominado c1 (presença do sítio de restrição) e o alelo mutante é denominado c2 (ausência do sítio de restrição). De acordo com as recomendações recentes para a nomenclatura de *CYP2E1*, o alelo \*1A equivale ao c1 e o alelo \*5 equivale ao c2 (<http://www.cypalleles.ki.se/cyp2e1.htm>).





O polimorfismo detectável pela enzima *Dra* I (7632T>A) está localizado no íntron 6 e gera a variante alélica *CYP2E1*\*6 (ausência do sítio de restrição - Figura 17). Outro polimorfismo localizado na região promotora de *CYP2E1* é uma inserção/deleção de 96 pb (alelos *CYP2E1*\*1D e \*1C) que pode regular a expressão desse gene. Alguns estudos têm demonstrado que as variantes alélicas *CYP2E1*\*5, \*6 e \*1D estariam associadas com um aumento da atividade enzimática (Watanabe *et al.*, 1994; Carriere *et al.*, 1996). Entretanto, outros autores não confirmaram qualquer relação desses polimorfismos com a atividade de *CYP2E1* (Powell *et al.*, 1998).

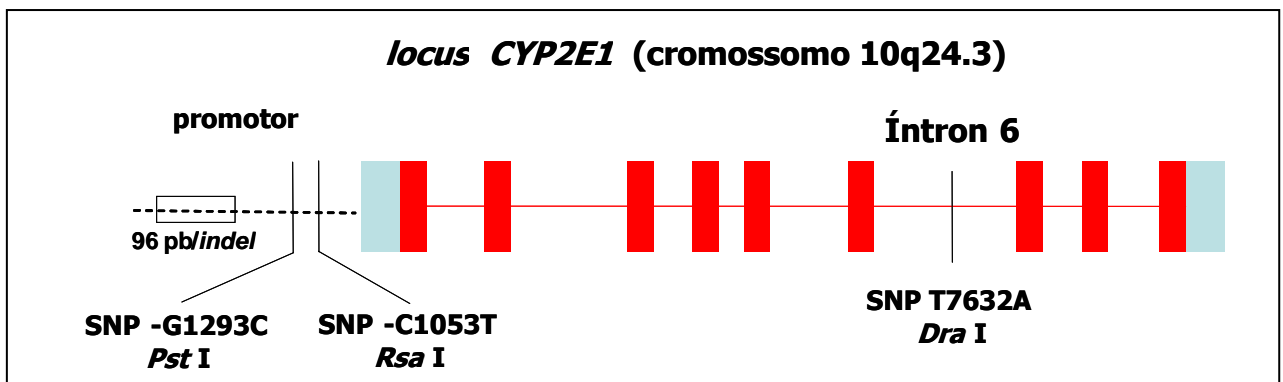


Figura 17: Localização dos polimorfismos no locus de *CYP2E1* e as respectivas enzimas de restrição utilizadas para detectar os diferentes genótipos (adaptado a partir de Das-Roy *et al.*, 2008).

Vários estudos têm descrito a influência dos polimorfismos em *CYP2E1* no desenvolvimento de vários tipos de câncer. Entretanto, os resultados são contraditórios. A maioria dos trabalhos destaca a mutação -1053C>T na região promotora do gene, a qual varia significativamente em termos de frequência, dependendo do grupo étnico estudado. O alelo mutante está presente com uma frequência de 2-8% em eurodescendentes enquanto que em asiáticos varia entre 25-36% (Neuhaus *et al.*, 2004). Em 2003, Huang e colaboradores mostraram a associação do genótipo \*1A/\*1A com o risco de desenvolvimento de lesão hepática induzida pela isoniazida, em pacientes adultos com tuberculose, independentemente do perfil de acetilação (OR 2,52; IC 95% 1,26-5,05 - Huang *et al.*, 2003). Posteriormente, Vuilleumier e colaboradores observaram a mesma associação entre o genótipo tipo-selvagem e a hepatotoxicidade induzida pela isoniazida, durante a quimioprofilaxia para a tuberculose (Vuilleumier *et al.*, 2006). Entretanto, um estudo com crianças indianas e outro com a população coreana não observaram nenhuma relação entre os efeitos adversos hepáticos com o genótipo \*1A/\*1A de *CYP2E1*, durante o tratamento anti-TB (Roy *et al.*, 2006; Cho *et al.*, 2007). Assim, novos estudos em diferentes populações e com um número amostral maior são necessários para verificar a real influência dos polimorfismos no gene *CYP2E1* com a ocorrência de lesão hepática durante o tratamento contra a tuberculose.



## 1.8) Genética molecular das glutatona S-transferases

As GSTs constituem uma superfamília de enzimas multifuncionais e ubíquas que desempenham um papel importante na desintoxicação celular, protegendo macromoléculas de ataques de reativos eletrofílicos. São enzimas de Fase II que catalisam o ataque nucleofílico da glutatona (GSH), em componentes que contêm um átomo eletrofílico de carbono, nitrogênio ou enxofre. A conjugação da GSH com esses compostos quase sempre leva a formação de produtos menos reativos e mais solúveis em água, prontos para a excreção pelo organismo (Hayes, *et al.*, 2005; Parl, 2005).

As glutatona S-transferases são de grande interesse para farmacologistas e toxicologistas, uma vez que são alvos de drogas no tratamento de asma e câncer além de metabolizarem drogas, inseticidas, herbicidas, carcinógenos e produtos de estresse oxidativo. Polimorfismos nos genes que codificam para as diferentes GSTs são frequentemente correlacionados com a suscetibilidade a diversos tipos de câncer, bem como à doença hepática alcoólica (Strange *et al.*, 2000; Strange *et al.*, 2001; Coughlin & Hall, 2002; Hayes, *et al.*, 2005; Parl, 2005).

Em humanos, oito famílias gênicas de GSTs solúveis (ou citossólicas) já foram descritas: alfa ( $\alpha$ ) localizada no cromossomo 6, mu ( $\mu$ ) no cromossomo 1; teta ( $\theta$ ) no cromossomo 22, pi ( $\pi$ ) no cromossomo 11; zeta ( $\zeta$ ) no cromossomo 14, sigma ( $\sigma$ ) no cromossomo 4; kappa ( $\kappa$ ) (localização cromossômica não determinada) e omega ( $\omega$ ) no cromossomo 10 (Strange *et al.*, 2001). Essa classificação é baseada na identidade entre sequências de aminoácidos, especificidade de substratos, afinidade química, estrutura protéica e comportamento cinético da enzima. Essas enzimas são altamente expressas no fígado constituindo até 4% do total de proteínas solúveis, mas podem ser observadas em diversos outros tecidos (Landi, 2000). As GSTs apresentam uma sobreposição de substratos específicos e a deficiência em uma isoforma pode ser compensada por outras isoformas. A glutatona S-transferase mu (GSTM1), glutatona S-transferase teta (GSTT1) e glutatona S-transferase pi (GSTP1) têm sido as isoformas mais estudadas (Pemble, *et al.*, 1994; Geisler & Olshan, 2001; Henrion-Caude, *et al.*, 2002; Vineis *et al.*, 2004; Raimondi *et al.*, 2006; Holley *et al.*, 2007).

A subfamília GST mu é codificada por cinco genes arranjados em *tandem* (5\_ GSTM4-GSTM2-GSTM1-GSTM5-GSTM3-\_3), formando um *cluster* gênico de 100 kb no cromossomo 1p13.3 (Figura 18). Já foram identificados polimorfismos em GSTM1 e as



consequências clínicas dos genótipos resultantes das combinações dos alelos *GSTM1\*0*, *GSTM1\*A* e *GSTM1\*B* têm sido amplamente investigadas (Brockmöller *et al.*, 1994; Cotton *et al.*, 2000; Coughlin & Hall, 2002; Parl, 2005). Os indivíduos que possuem o genótipo homozigoto nulo para *GSTM1* (*GSTM1\*0/GSTM1\*0*) não expressam essa proteína. Desta maneira, a ausência deste gene pode causar um maior acúmulo de metabólitos reativos no organismo, aumentando a probabilidade de interação com as macromoléculas celulares e a iniciação do processo tumoral. *GSTM1\*A* e *GSTM1\*B* diferem apenas em uma base no éxon 7 e codificam monômeros que formam dímeros ativos. A atividade catalítica das enzimas codificadas por ambos os alelos são similares (Widersten *et al.*, 1991). Em 1997, McLellan e colaboradores descreveram uma variante alélica em dois indivíduos da Arábia Saudita que demonstraram uma atividade enzimática ultrarrápida devido à presença de duas cópias de *GSTM1* entre os genes *GSTM2* e *GSTM5* (McLellan *et al.*, 1997).

O gene *GSTM1* é flanqueado por duas regiões quase idênticas de 4,2 Kb. A deleção *GSTM1\*0*, causada por uma recombinação homóloga entre essas duas regiões repetidas, resulta na perda de um fragmento de 16 Kb contendo o gene *GSTM1* inteiro (Figura 18). *GSTM1* é excisado de forma precisa deixando os genes adjacentes, *GSTM2* e *GSTM5*, intactos (Parl, 2005). Em estudo de espécimes de fígado de cento e sessenta e oito indivíduos japoneses autopsiados, Harada e colaboradores (1987) observaram que o alelo nulo *GSTM1\*0* era mais frequente nos fígados portadores de hepatite e hepatocarcinoma em relação a fígados controles (Harada *et al.*, 1987).

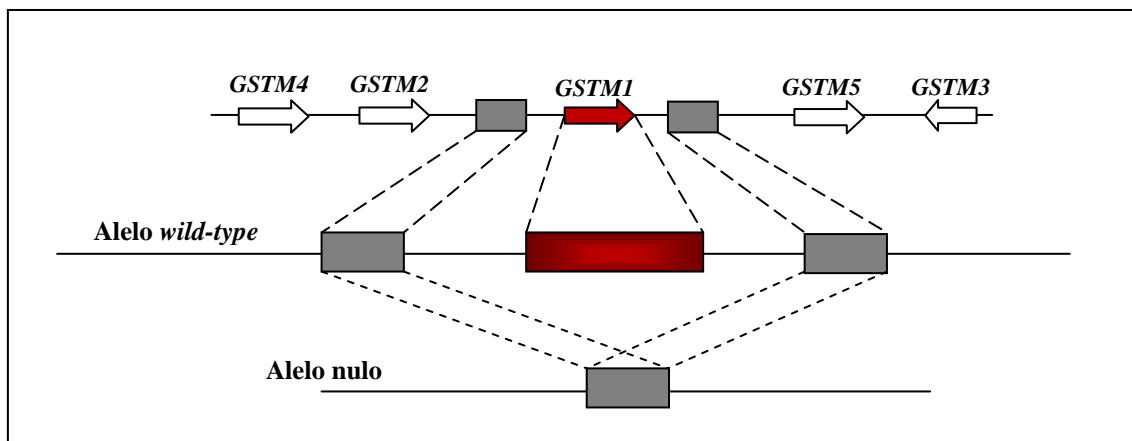


Figura 18: Localização estrutural do *cluster* gênico de 100 kb que codifica para a subfamília GST mu (cromossomo 1p13.3). A figura indica o evento de recombinação homóloga que pode acontecer nesse *locus* e originar o alelo nulo (*GSTM1\*0* – ausência de *GSTM1*). Figura adaptada de Parl, 2005.

A subfamília GST teta consiste de dois genes, *GSTT1* e *GSTT2*, localizados no cromossomo 22q11.2 e separados por aproximadamente 50 Kb (Figura 19). Uma análise da porção de 119 Kb contendo os dois genes revelou duas regiões flanqueando *GSTT1*, HA3 e



HA5, com mais de 90% de homologia. HA3 e HA5 contêm duas regiões de 403 pb idênticas e, de forma semelhante a *GSTM1\*0*, a deleção de *GSTT1\*0* é causada provavelmente por recombinação homóloga entre essas duas regiões repetidas (Figura 19) (Parl, 2005).

Diferentemente de *GSTM1*, *GSTT1* apresenta as atividades de desintoxicação e ativação. Por exemplo, *GSTT1* é capaz de ativar di-halometanos a diclorometanos, aparentemente envolvidos no desenvolvimento de câncer de fígado e pulmão em camundongos (Pemble *et al.*, 1994). Em humanos, *GSTT1* é expressa também nos eritrócitos e possivelmente desempenha um papel mais global na desintoxicação inicial de xenobióticos e carcinógenos.

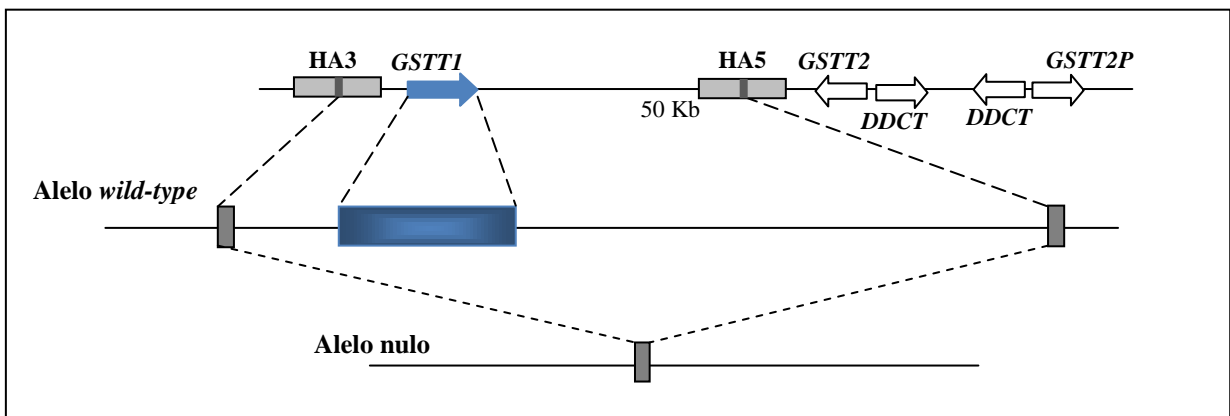


Figura 19: Localização estrutural do *cluster* gênico que codifica para a subfamília GST teta (cromossomo 22q11.2). A figura mostra duas regiões repetitivas que flanqueiam o gene *GSTT1* e que, através de recombinação homóloga, permitem a sua excisão do cromossomo (alelo *GSTT1\*0*). Figura adaptada de Parl, 2005.

Deficiências na atividade de GST, devidas aos genótipos nulos de *GSTM1* e *GSTT1*, podem modular a suscetibilidade ao desenvolvimento de hepatotoxicidade induzida por drogas e xenobióticos. Além disso, foi observado que as frequências dos alelos *GSTM1\*0* e *GSTT1\*0* variam de acordo com os diferentes grupos étnicos (Landi, 2000; Parl, 2005). A lesão hepática induzida pela isoniazida tem sido associada à depleção do conteúdo de glutatona, com uma redução da atividade de GST, em modelo animal com hepatotoxicidade por drogas anti-TB (Sodhi *et al.*, 1996).

Em 2001, Roy e colaboradores demonstraram que indivíduos homozigotos nulos para *GSTM1* apresentavam um risco relativo de 2,12 para desenvolver hepatotoxicidade induzida por drogas anti-TB. Entretanto, esses autores não verificaram a associação do genótipo nulo de *GSTT1* com o desenvolvimento de reação adversa a isoniazida (Roy *et al.*, 2001). De forma similar, em outro estudo realizado na população tailandesa, também foi verificado que somente o genótipo nulo de *GSTM1* aumenta o risco de desenvolvimento de lesão hepática (OR 2,23; IC 95% 1,07-4,67 - Huang *et al.*, 2007). Por sua vez, Leiro e colaboradores observaram o contrário: indivíduos com genótipo nulo de *GSTT1* apresentaram um risco



aumentado em desenvolver hepatotoxicidade induzida por drogas anti-TB e nenhuma associação significativa foi observada entre o genótipo *GSTM1*\*0/\*0 e a lesão hepática em pacientes sob tratamento anti-TB (Leiro *et al.*, 2008). Esses estudos sugerem um efeito protetor das glutatona S-transferases em relação aos efeitos hepatotóxicos da isoniazida; entretanto, novas populações devem ser avaliadas para verificar quais desses polimorfismos podem ser utilizados como marcadores genéticos para o risco de ocorrência de efeitos colaterais durante tratamento anti-TB.

Os brasileiros formam uma das populações mais heterogêneas do mundo, resultado de mais de cinco séculos de cruzamento interétnico: europeus, representados principalmente por portugueses, escravos africanos e ameríndios. Quando os portugueses chegaram ao Brasil, em 22 de abril de 1500, havia aproximadamente 2,5 milhões de indígenas vivendo nesta área (IBGE, 2000). A primeira colonização neste país envolveu principalmente integrantes do sexo masculino, o que acabou ocasionando o casamento entre homens europeus com mulheres indígenas. Posteriormente, com a chegada dos escravos durante o ciclo econômico da cana-de-açúcar, começaram a ocorrer relacionamentos entre europeus e africanas. Após a abertura dos portos para outras nações, o Brasil recebeu ainda um grande número de imigrantes de várias partes da Europa, incluindo Itália, Espanha e Alemanha e, no século 20, foi a vez da imigração asiática principalmente por parte dos japoneses, libaneses e sírios contribuindo ainda mais para a miscigenação brasileira (IBGE, 2000).

O alto nível de diversidade genômica e a extensa miscigenação da população brasileira têm implicações importantes nos estudos de farmacogenética, uma vez que não é possível a extrapolação dos dados encontrados em grupos étnicos fechados para a maioria das pesquisas realizadas no país. Além disso, a contribuição genética de cada etnia parental aconteceu de forma diferenciada ao longo das regiões do país fazendo com que cada estado brasileiro apresente a sua particularidade no perfil de variantes gênicas. Sendo assim, um estudo de associação entre os polimorfismos de genes metabolizadores de drogas e a ocorrência de hepatotoxicidade seria de grande relevância na população brasileira já que a tuberculose é uma doença com alta incidência no Brasil, onde a quimioprofilaxia com INH é preconizada, e levando-se em conta que a distribuição dos alelos de *NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* varia de acordo com a etnia.



## ***2- OBJETIVOS***



**Geral:** Avaliar a possível associação entre as variantes alélicas dos genes *NAT2*, *CYP2E1*, *GSTM1* e *GSTT1* e a ocorrência de efeitos adversos hepáticos em pacientes sob tratamento anti-tuberculose, com esquema contendo isoniazida, atendidos em um hospital geral do Rio de Janeiro, referência para AIDS.

**Objetivos específicos:**

- Determinar a frequência de SNPs no gene *NAT2* em amostras de duas regiões brasileiras (Rio de Janeiro e Goiás);
- Identificar novos SNPs na região codificante de *NAT2* e caracterizar novos alelos/haplótipos na população estudada;
- Determinar os haplótipos de *NAT2* mais frequentes e os perfis de acetilação presentes nas populações de estudo, comparando as frequências alélicas encontradas no Rio de Janeiro e em Goiás;
- Avaliar os efeitos de polimorfismos importantes na estrutura molecular tridimensional da proteína *NAT2*;
- Identificar a frequência dos alelos c1 (-1053C) e c2 (-1053T) de *CYP2E1* no grupo de estudo proveniente do RJ;
- Identificar a frequência dos genótipos nulos para os genes *GSTM1* e *GSTT1* no grupo de estudo proveniente do RJ.
- Avaliar a associação entre o fenótipo de acetilação lenta (presença de dois alelos de *NAT2* que caracterizam baixa atividade) com a hepatotoxicidade e/ou hepatite medicamentosa em pacientes sob tratamento anti-TB, com esquema contendo isoniazida, residentes no Rio de Janeiro;
- Avaliar a possível associação entre a ocorrência de hepatotoxicidade e/ou hepatite medicamentosa e a presença de polimorfismos no gene que codifica para *CYP2E1* em pacientes sob tratamento anti-TB, com esquema contendo isoniazida, residentes no Rio de Janeiro;
- Avaliar a possível associação entre a ocorrência de hepatotoxicidade e/ou hepatite medicamentosa e a presença de um ou dois alelos nulos das glutatona S transferases M1 e T1 em pacientes sob tratamento anti-TB, com esquema contendo isoniazida, residentes no Rio de Janeiro.



**Metas:**

- Obter informações a respeito da variabilidade dos genes em estudo tendo em vista sua grande relevância na farmacologia e a existência de poucos dados disponíveis na população brasileira;
- Descrever as frequências dos diferentes polimorfismos nos genes que codificam para NAT2, CYP2E1 e GSTs na população estudada;
- Fornecer dados que serão úteis para a avaliação futura de novas drogas, modificação da conduta terapêutica ou correção da dosagem individual com base nos perfis genômicos envolvidos, tendo em vista que a concentração de isoniazida recomendada em nosso país (400mg/dia) difere da preconizada pela OMS (300mg/dia).

**Hipótese nula:**

- Pacientes com tuberculose em tratamento específico, contendo isoniazida, e que apresentam polimorfismos genéticos no gene que codifica para NAT2, resultando num fenótipo de acetiladores lentos, bem como nos genes *CYP2E1*, *GSTM1* e *GSTT1* apresentam um risco de desenvolverem hepatotoxicidade similar aos acetiladores rápidos e aqueles não possuem os polimorfismos nos outros genes em questão.

**Hipótese alternativa:**

- Pacientes com tuberculose em tratamento específico, contendo isoniazida, e que apresentam polimorfismos genéticos no gene que codifica para NAT2, resultando num fenótipo de acetiladores lentos, bem como nos genes *CYP2E1*, *GSTM1* e *GSTT1* apresentam risco maior de desenvolverem hepatotoxicidade e/ou hepatite medicamentosa.





### ***3- RESULTADOS***



## CAPÍTULO I

Essa etapa incluiu a padronização da amplificação e seqüenciamento da região codificante de *NAT2*, a análise da distribuição dos alelos de *NAT2* e o perfil de acetilação de indivíduos residentes do Rio de Janeiro e de Goiás. Esses resultados encontram-se **publicados** no manuscrito:

“Genetic profile of the Arylamine N-Acetyltransferase 2 coding gene among individuals from two different regions of Brazil”.

Revista de publicação: *Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis* 2007; 624(1-2):31-40.

## Genetic profile of the arylamine *N*-acetyltransferase 2 coding gene among individuals from two different regions of Brazil

Raquel L.F. Teixeira<sup>a</sup>, Antonio B. Miranda<sup>b</sup>, Antonio G. Pacheco<sup>c</sup>,  
Márcia Q.P. Lopes<sup>a</sup>, Joseane Fonseca-Costa<sup>d</sup>, Marcelo F. Rabahi<sup>e</sup>,  
Hedi M. Melo<sup>f</sup>, Afrânio L. Kritski<sup>d</sup>, Fernanda C.Q. Mello<sup>d</sup>,  
Philip N. Suffys<sup>a</sup>, Adalberto R. Santos<sup>a,\*</sup>

<sup>a</sup> *Laboratory of Molecular Biology Applied to Mycobacteria, Department of Mycobacteriosis, IOC/Fiocruz, Avenida Brasil 4365, Manguinhos, Rio de Janeiro, CEP 21040-900, Brazil*

<sup>b</sup> *Laboratory for Functional Genomic and Bioinformatics, Department of Biochemistry and Molecular Biology, Fiocruz, Avenida Brasil 4365, Rio de Janeiro, CEP 21040-900, Brazil*

<sup>c</sup> *Department of Epidemiology and Quantitative Methods in Health, National School of Public Health/Fiocruz, Rua Leopoldo Bulhões 1.480, Manguinhos, Rio de Janeiro, CEP 21040-900, Brazil*

<sup>d</sup> *Hospital Complex - HUCFF-IDT, Federal University of Rio de Janeiro, Rua Professor Rodolpho Paulo Rocco 255, Ilha do Fundão, Rio de Janeiro, CEP 21941-913, Brazil*

<sup>e</sup> *Anuar Auad Infectious Disease Reference Hospital, Av.do Contorno, s/n, Goiânia, CEP 74853-400, Brazil*

<sup>f</sup> *Santa Maria State Hospital, Estrada do Rio Pequeno 656, Rio de Janeiro, CEP 22723-195, Brazil*

Received 11 July 2006; received in revised form 11 January 2007; accepted 1 March 2007

Available online 19 April 2007

### Abstract

Arylamine *N*-acetyltransferase 2 is the main enzyme responsible for the isoniazid metabolization into hepatotoxic intermediates and the degree of hepatotoxicity severity has been attributed to genetic variability in the *NAT2* gene. The main goal of this study was to describe the genetic profile of the *NAT2* gene in individuals from two different regions of Brazil: Rio de Janeiro and Goiás States. Therefore, after preparation of DNA samples from 404 individuals, genotyping of the coding region of *NAT2* was performed by direct PCR sequencing. Thirteen previously described SNPs were detected in these Brazilian populations, from which seven: 191 G>A; 282 C>T; 341 T>C; 481 C>T; 590 G>A; 803 A>G and 857 G>A are the most frequent in other populations. The presence of so-called ethnic-specific SNPs in our population is in accordance with the Brazilians' multiple ancestry. Upon allele and genotype analysis, the most frequent *NAT2* alleles were respectively *NAT2\*5B* (33%), *NAT2\*6A* (26%) and *NAT2\*4* (20%) being *NAT2\*5/\*5* the more prevalent genotype (31.7%). These results clearly demonstrate the predominance in the studied Brazilian groups of *NAT2* alleles associated with slow over the fast and intermediate acetylator genotypes. Additionally, in Rio de Janeiro, a significantly higher frequency of intermediate acetylation status was found when compared to Goiás (42.5% versus 25%) ( $p=0.05$ ), demonstrating that different regions of a country with a population characterized by a multi-ethnic ancestry may present a large degree of variability in *NAT2* allelic frequencies. This finding has implications in the determination of nationwide policies for use of appropriate anti-TB drugs.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** *NAT2*; Polymorphism; Genotyping; Sequencing; Admixed population; Brazilian subjects

\* Corresponding author. Tel.: +55 21 2598 4289; fax: +55 21 2270 9997.

E-mail address: [adalbertorezende@yahoo.com.br](mailto:adalbertorezende@yahoo.com.br) (A.R. Santos).

## 1. Introduction

The increasing incidence of tuberculosis (TB) worldwide together with the spreading of multidrug resistant strains (MDR) of *Mycobacterium tuberculosis* are the main reasons why this disease is considered an important public health problem. In addition, control of this endemic disease is also hampered by liver adverse reactions (ADRs) induced by anti-TB drugs that are important causes of morbidity/mortality [1,2] and represent an annual cost of approximately 100 billion US\$ in the US [3].

Isoniazid (INH), one of the most important drugs used in anti-TB treatment and in TB chemoprophylaxis, is also the major drug involved in hepatotoxicity [4,5]. Arylamine *N*-acetyltransferase 2 (NAT2), a cytosolic phase II conjugation enzyme, is responsible for the *N*-acetylation of arylamine and hydrazine xenobiotics, including drugs and carcinogens being consequently associated with drug effects or toxicity levels. In humans, two different genes (*NAT1* and *NAT2*) located in the chromosome 8 p22 encode *N*-acetyltransferases with partly overlapping substrate specificities [6–8]. *N*-Acetyltransferase 2 is mainly responsible for INH metabolism intermediating the generation of hepatotoxins. Both loci are highly polymorphic and differences in INH-induced toxicity, discovered over 40 years ago, have been attributed to genetic variability in the *NAT2* gene [9–11].

Polymorphism at the *N*-acetylation level is an example of a pharmacogenetic trait that implies on interindividual and interethnic differences in response to xenobiotics. From the clinical point of view, *N*-acetylation polymorphism has been implicated with therapeutic effects and adverse reactions of various amine-containing drugs and in the susceptibility to cancers and others disorders. In fact, epidemiological studies have suggested that several *NAT1* and *NAT2* alleles are linked to increased susceptibility to drug toxicity and cancer risk caused by arylamine exposure [5,7,12–16].

Variations within the *NAT2* gene were initially studied because of their association with INH metabolism and drug-induced side effects in TB patients [17,18]. The wild type *NAT2*\*4 allele and 35 other variants of *NAT2* gene were identified and classified in human populations depending on the haplotype determined by the combination of up to four out of the 17 different single nucleotide polymorphisms (SNPs) present throughout the *NAT2* coding region ([www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html)). Several other rare SNPs were described in other populations and reported (<http://www.ncbi.nlm.nih.gov/GenBank>), but for allele

determination they are not taken into account. Considering only the seven most common SNPs, four result in an amino acid change leading to a significant decrease in acetylation capacity (SNPs 191 G>A, 341 T>C, 590 G>A and 857 G>A). The other three (282 C>T, 481 C>T and 803 A>G) are silent mutations or do not alter the phenotype [19].

It has been demonstrated that the frequencies of genotypes and phenotypes in different ethnic closed populations are variable [20]. Ethnic-specific SNPs like 191 G>A, which is commonly found in the *NAT2*\*14 haplotype cluster and is associated with low INH acetylation, is present in African Americans and Sub-Saharan Africans, but it is absent or rarely found in European descents [6,21,22] where the most frequent SNPs are the 341 T>C, and the frequently associated 481 C>T and 282 C>T. In this population the frequency of defective alleles lacking these SNPs is less than 3% [23,24]. Another variant, 857 G>A, although relatively uncommon in most of ethnic closed populations, is present in a high frequency in Asians and Amerindians [25–28]. In mixed Hispanic population, a high impact of the genetic admixture of Asian origin (Central American Indians) and European descents (Spaniards) was reflected by the high heterogeneity of genotypes and phenotypes observed [29].

Brazilians are members of one of the most heterogeneous population in the world, result of five centuries of ethnic admixture of people from three continents: the European colonizers mainly represented by the Portuguese, the African slaves and the autochthonous Amerindians [30]. Thus, the admixture of ethnicities is spread all over the country with predominance of certain groups depending on the region. Having in mind the diversity of *NAT2* allele distribution related to this factor it is reasonable to expect different *N*-acetylation statuses in these different regions. However, INH-containing schemes for TB-treatment in Brazil use a standardized dosage for INH administration irrespectively of the region.

Recently, the genetic diversity of *NAT2* has been clearly demonstrated among thirteen different native Americans populations including three from Brazil (Karitiana, Surui and Xavante) [31]. We described here the first genetic profile of the *NAT2* gene of individuals from two distinct urban regions of Brazil: Rio de Janeiro and Goiás, respectively in the Southeast and Center West region. These two States reflect a distinct population behavior which results in different ethnic characteristics. Rio de Janeiro is one of the most important economic and touristic poles of the country and its cosmopolite characteristic results in high national

and international migration rates. On the other side, in Goiás, this migratory flux is not observed with favors a more homogeneous population with predominance of Amerindians and Amerindians/European descents.

## 2. Methods

### 2.1. Selection of subjects and sample collection

Four hundred and four consanguineously unrelated individuals from two different geographic regions of Brazil were enrolled in this study. They comprised 248 active TB patients and 50 close healthy contacts of smear positive pulmonary TB cases from the University Hospital Complex: Thoracic Institute/Clementino Fraga University Hospital from Federal University of Rio de Janeiro (UFRJ) and Reference Santa Maria State Hospital, both situated in Rio de Janeiro State and 106 asymptomatic Health Care Workers (HCWs) from Anuar Auad Infectious Disease Reference Hospital of the State Health Secretariat of Goiás State, Brazil. This study was approved by the Institutional Review Boards of the University Hospital Clementino Fraga Filho (UFRJ), of the State Health Secretariat of Goiás State, the Oswaldo Cruz Foundation and the Brazilian Ministry of Health. All participants were informed about the objectives of the project and a written informed consent was obtained.

Because of the complete lack of information regarding this issue in Brazil, sample size was estimated after a number of assumptions. We worked with a 3:1 ratio. Power for two-sample Binomial tests was calculated for different sample sizes in order to establish a convenient one that would give us at least 80% of power. A review of the literature revealed that for these SNPs, varying frequencies in one of the alleles is present, so we also tested frequencies ranging from 0.01 to 0.4. Given these conditions and a type I error of 0.05 we agreed that a sample of around 400 individuals would give us enough power to detect as low as an odds ratio of 1.9 for the 0.4 frequency group (i.e. a frequency of 0.56 in the other group) up to as high as 7 for the 0.01 frequency group (0.066 in the other). These figures seemed reasonable for us, given that for small frequencies in one group, larger differences should be reached in order to provide a clinical/biological significant difference.

A volume of 3 ml of venous blood was collected from each volunteer and stored at  $-20^{\circ}\text{C}$ . Genomic DNA was isolated from 100  $\mu\text{l}$  of frozen whole blood using the FlexiGene DNA Kit (Qiagen Inc., USA), according to the manufacturer's specifications. After extraction, DNA samples were stored at  $-20^{\circ}\text{C}$ .

### 2.2. NAT2 genotyping

Two sets of primers for PCR amplification and/or sequencing (NAT2 EF 5' TTA GTC ACA CGA GGA AAT CAA A 3' and NAT2 ER 5' AAA TGC TGA CAT TTT TAT GGA TGA 3'. NAT2 IF 5' ACC ATT GAC GGC AGG AAT TA 3' and NAT2 IR 5' TGG TCC AGG TAC CAG ATT CC 3') were designed using the Primer3 software (<http://frodo.wi.mit.edu/cgi-bin/>

[primer3/primer3.www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.www.cgi)). Genotyping of the NAT2 coding region was performed by generation of the intronless 1093 bp PCR fragment and direct sequencing. For PCR amplification 100 ng of genomic DNA were added to a reaction mixture containing 200 ng of each primer (NAT2EF and NAT2ER), 0.2 mM of each dNTP, 4.0 mM  $\text{MgCl}_2$ , 10% glycerol, 10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% Triton X-100, and 1.1 U of DNA polymerases (mixture of 1 U of *Taq* DNA polymerase and 0.1 U of *Tli* DNA polymerase - Promega Corporation, USA). Samples were incubated at  $94^{\circ}\text{C}$  for 5 min, followed by 35 cycles  $94^{\circ}\text{C}$  for 1 min,  $57.5^{\circ}\text{C}$  for 1 min and  $72^{\circ}\text{C}$  for 1 min; final extension took place at  $72^{\circ}\text{C}$  for 5 min. Evaluation of the PCR product was achieved by electrophoresis on a 1% agarose gel followed by ethidium bromide (0.5  $\mu\text{g}/\text{ml}$ ) staining. The PCR products were purified with ammonium acetate 2.5 M and used for direct sequencing. Four different sequencing reactions, using the described primers, were performed separately for each sample. Sequencing was performed using ABI PRISM Big Dye Terminator v. 3.1 Kit (PE Applied BioSystems), according to the manufacturer's recommendations, on an ABI PRISM 3730 DNA Analyser (PE Applied BioSystems).

### 2.3. Computational analysis

Sequence data of each sample were analyzed for SNPs identification after alignment with the reference sequence AY331807 (<http://www.ncbi.nlm.nih.gov/GenBank>) through SeqScape v.2.5 software (Applied BioSystems-<http://www.appliedbiosystems.com>).

### 2.4. Statistical analysis

Expected genotype frequencies were calculated from respective single allele frequencies and were consistent with Hardy Weinberg equilibrium. Pair-wise linkage disequilibrium was also tested for the loci studied using conventional methods as implemented in the R package "genetics" [32].

Overall linkage disequilibrium, estimation of haplotypes with unknown phase via the EM algorithm and association of haplotype distributions with geographical region, either using a score statistic or generalized linear models were performed as implemented by the R package "haplo.stats" [33], originally proposed by Schaid et al. [34]. For genotype comparisons, log-linear models were employed and comparisons of frequencies with published data were performed by Fisher's exact test, with Bonferroni correction. All the analyses were carried out with the statistical package R [35].

## 3. Results

### 3.1. SNPs profile of the human NAT2 gene among Brazilians subjects

Using the direct PCR sequencing approach, analysis of the 404 DNA samples enrolled in this study revealed

the presence of 13 previously described SNPs of which 7 represent the most frequently found in different populations and are registered in the official site of *NAT2* nomenclature ([www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html)). Five of the remaining SNPs have been recently reported in sub-Saharan African populations [36,37] and one (345 C>T) is deposited in genetic databases (GenBank-Entrez SNP). The allele frequencies of the 13 SNPs found in the present study and their presumed influence on the protein sequence are shown in Table 1. No significant differences were found between the frequencies of any SNP when Rio de Janeiro and Goiás were compared (Table 1).

### 3.2. Determination of the *NAT2* allele variants present in Rio de Janeiro and Goiás

A linkage disequilibrium test (LDT) was carried out with the identified SNPs and it was observed that four SNPs (345 C>T; 403 C>G; 683 C>T and 838 G>A) were not in LD with any other being consequently excluded for allele characterization.

Sixteen different haplotypes were identified in the study population by the use of the Score Test (R for Windows Version 2.2.1) [35] allowing the characterization of the corresponding *NAT2* alleles. Table 2 shows the frequency of each allele variant found in the overall studied population and after stratification in the two different studied regions. Among the functional alleles, the *NAT2*\*4 (wild type) was the most frequent with an overall frequency of 20% whereas the alleles \*5B and \*6A were the more frequent (33% and 26% respectively)

among the defective alleles (those associated with slow acetylation). Comparison of the frequencies between Rio de Janeiro and Goiás by the use of Generalized Linear Model (GLM) revealed significant differences in the distribution of the *NAT2*\*4. (*p*-value 0.02), which were 1.6 times higher in Goiás. Additionally, Table 2 shows the lower allele diversity found in Goiás compared to Rio de Janeiro. Out of the 16 allele variants characterized, 15 were identified in Rio de Janeiro and only 11 in Goiás.

### 3.3. Genotype characterization of the *NAT2* gene and preliminary determination of the acetylation status

For genotype characterization, a total of six SNPs were excluded (two synonymous and four that were not in LD with any other) since they were not informative for phenotype prediction. From the remaining samples only those homozygous for all SNPs found or heterozygous for only one SNP were considered. In this way, genotype determination was achieved for 50% of the samples (202 out of 404).

The distribution of the genotype frequencies and the predicted phenotypes in Rio de Janeiro and Goiás is shown in Table 3. A total of 14 different genotypes was observed with an overall predominance of the low acetylation status (49%) being the \*5/\*5 genotype the most frequent (31.7%). The characterization of the fast and intermediate phenotypes was predominantly determined respectively by the wild type \*4/\*4 and \*4/\*6 genotypes while the slow acetylation phenotype was determined by both \*5/\*5 and \*6/\*6 genotypes. Out

Table 1  
Frequency of point mutations in the *N-acetyltransferase (NAT2)* gene among Brazilian subjects from Rio de Janeiro and Goiás

SNPs	Total		Rio de Janeiro		Goiás		<i>p</i> -value	Effect	aa change
	<i>n</i> <sup>a</sup>	%	<i>n</i> <sup>b</sup>	%	<i>n</i> <sup>c</sup>	%			
70A	01	0.1	1	0.2	0	0	0.8416	Missense	L24I
191A <sup>d</sup>	33	4.1	28	5	5	2	0.3055	Missense	R64Q
282T <sup>d</sup>	298	37	226	38	72	34	0.4686	Silent	None
341C <sup>d</sup>	306	38	236	40	70	33	0.2312	Missense	I114T
345T	02	0.3	0	0	2	1	0.7991	Silent	None
403G	05	0.6	4	1	1	0.5	0.823	Missense	L135V
481T <sup>d</sup>	284	35	217	36	67	32	0.3739	Silent	None
590A <sup>a</sup>	215	27	155	26	60	28	0.6462	Missense	R197Q
609T	03	0.4	2	0.3	1	0.5	0.8439	Missense	E203D
683T	02	0.2	2	0.3	0	0	0.8337	Missense	P228L
803G <sup>d</sup>	322	40	245	41	77	36	0.3877	Missense	K268R
838A	04	0.5	3	0.5	1	0.5	0.9682	Missense	V280M
857A <sup>d</sup>	36	4.4	27	4.5	9	4.2	0.9028	Missense	G286E

<sup>a</sup> Frequency refers to 808 alleles.

<sup>b</sup> Frequency refers to 596 alleles.

<sup>c</sup> Frequency refers to 212 alleles.

<sup>d</sup> Seven most common SNPs in the world.

Table 2  
Frequency of the *NAT2* allele variants among individuals from Rio de Janeiro and Goiás

<i>NAT2</i> alleles	T70A	G191A	C282T	T341C	C481T	G590A	G609T	A803G	G857A	Total	RJ <sup>a</sup>	GO <sup>b</sup>	OR <sup>c</sup>	95% CI	<i>p</i> -value
4	T	G	C	T	C	G	G	A	G	0.20	0.17	0.28	<b>1.71</b>	<b>1.11–2.64</b>	<b>0.02</b>
12A	T	G	C	T	C	G	G	<b>G</b>	G	0.04	0.04	0.03	0.91	0.39–2.11	0.83
13	T	G	<b>T</b>	T	C	G	G	A	G	0.03	0.03	<0.01	0.16	0.02–1.21	0.08
5A	T	G	C	<b>C</b>	<b>T</b>	G	G	A	G	0.02	0.03	<0.01	0.17	0.02–1.36	0.10
5B	T	G	C	<b>C</b>	<b>T</b>	G	G	<b>G</b>	G	0.33	0.34	0.31	–	–	–
5C	T	G	C	<b>C</b>	C	G	G	<b>G</b>	G	0.03	0.03	0.01	0.54	0.15–1.86	0.33
6A	T	G	<b>T</b>	T	C	<b>A</b>	G	A	G	0.26	0.26	0.28	1.14	0.76–1.71	0.53
6B <sup>d</sup>	T	G	C	T	C	<b>A</b>	G	A	G	<0.01	<0.01	0	–	–	–
7B	T	G	<b>T</b>	T	C	G	G	A	<b>A</b>	0.04	0.04	0.04	0.89	0.38–2.11	0.80
14A <sup>d</sup>	T	<b>A</b>	C	T	C	G	G	A	G	<0.01	<0.01	0.01	–	–	–
14B	T	<b>A</b>	<b>T</b>	T	C	G	G	A	G	0.03	0.04	<0.01	0.24	0.05–1.06	0.06
NC <sup>d</sup>	T	G	C	T	C	G	<b>T</b>	A	G	<0.01	<0.01	0	–	–	–
NC <sup>d</sup>	T	G	C	T	C	G	<b>T</b>	<b>G</b>	G	<0.01	<0.01	0	–	–	–
NC <sup>d</sup>	<b>A</b>	<b>A</b>	C	T	C	G	G	A	G	<0.01	<0.01	0	–	–	–
NC <sup>d</sup>	T	<b>A</b>	<b>T</b>	T	C	G	G	A	<b>A</b>	<0.01	<0.01	0	–	–	–
NC <sup>d</sup>	T	G	<b>T</b>	T	C	G	<b>T</b>	<b>G</b>	<b>A</b>	<0.01	0	<0.01	–	–	–

NC: not classified according to the official *NAT2* nomenclature. All estimatives and comparison were carried out according to Schaid et al. [34] and Sinnwell and Schaid [33].

<sup>a</sup> Rio de Janeiro (596 alleles).

<sup>b</sup> Goiás (212 alleles).

<sup>c</sup> Odds ratio.

<sup>d</sup> Rare alleles.

Table 3

Genotype characterization of *NAT2* gene and preliminary determination of acetylation status in 202 Brazilians subjects from two different regions

Cluster designation	Total		Rio de Janeiro		Goiás		<i>p</i> -value
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
*4/*4	16	7.9	7	4.8	9	16.1	<b>0.002</b>
*4/*12 <sup>a</sup>	5	2.5	4	2.7	1	1.8	–
*4/*13 <sup>a</sup>	2	1	2	1.4	0	0	–
*12/*12 <sup>a</sup>	3	1.5	2	1.4	1	1.8	–
*13/*13 <sup>a</sup>	1	0.5	1	0.7	0	0	–
Fast acetylation	27	<b>13.4</b>	16	<b>10.9</b>	11	<b>19.6</b>	–
*4/*5 <sup>a</sup>	3	1.5	3	2.1	0	0	–
*4/*6	46	22.8	36	24.7	10	17.9	0.37
*4/*7 <sup>a</sup>	6	3	4	2.7	2	3.6	–
*4/*14 <sup>a</sup>	4	2	4	2.7	0	0	–
*12/*5 <sup>a</sup>	9	4.5	7	4.9	2	3.6	–
*13/*6 <sup>a</sup>	8	4	8	5.5	0	0	–
Intermediate acetylation	76	<b>37.6</b>	62	<b>42.5</b>	14	<b>25</b>	–
*5/*5	64	31.7	49	33.6	15	26.8	0.25
*6/*6	34	16.8	18	12.3	16	28.6	<b>0.003</b>
*14/*14 <sup>a</sup>	1	0.5	1	0.7	0	0	–
Slow acetylation	99	<b>49</b>	68	<b>46.6</b>	31	<b>55.4</b>	–
Total	202	100	146	100	56	100	

<sup>a</sup> Genotype frequencies lower than 5%.

of the 14 genotypes identified, 6 were not found in Goiás. A comparison of the predicted phenotypes of Rio de Janeiro and Goiás showed an statistically significant higher frequency of the intermediate acetylation phenotype in Rio de Janeiro ( $p=0.05$ ). The same comparison for fast and slow acetylation phenotypes showed no statistic differences (Table 3). After clustering the genotypes with frequencies lower than 5% in a group designated as rare, a statistically significant difference was observed in the frequencies of the \*4/\*4 and \*6/\*6 genotypes upon comparison of Rio de Janeiro and Goiás ( $p$ -values 0.002 and 0.003, respectively). Furthermore, Table 3 shows that \*4/\*4 and \*6/\*6 genotypes were 3.4 and 2.3 times more frequent in Goiás than in Rio de Janeiro.

#### 3.4. Interethnic distribution of SNPs of *NAT2* gene

The sequencing-based determination of SNPs in the *NAT2* gene of Brazilians allowed, for the seven most prevalent SNPs worldwide, the comparison of the allele frequencies observed in our study with the frequencies observed in different ethnic populations around the world (Table 4). The SNP 191 G>A, mainly identified in African populations, and 857 G>A, of Asian origin but also prevalent in Amerindian populations, were found in our population. The frequency of 191 G>A was significantly higher ( $p<0.01$ ) in our population than

that of European descents, Germans and Polish and significantly lower ( $p<0.01$ ) than Sub-Saharan Africans (Guinea Bissau). The frequency of 857 G>A polymorphism, was significantly higher in Brazil than in European descents, Germans and Spanish ( $p<0.05$ ) and significantly lower than Thai, Chinese and Amerindians ( $p<0.01$ ). Additionally, the SNPs 341 T>C, 481 C>T and 803 A>G, all highly prevalent in European and African populations, were also frequently observed in Brazil (Table 4). The frequency of the SNP 341 T>C in Brazilians was significantly higher than in Japanese, Chinese and Amerindians populations ( $p<0.01$ ). The SNP 481 C>T was also more frequent in our studied population than in Asian populations like Japanese and Chinese ( $p<0.001$ ) as well as in Amerindians ( $p<0.001$ ) but less frequent than in European descents ( $p<0.05$ ), while SNP 803 A>G was observed at a higher frequency than in Japanese and Amerindians ( $p<0.001$ ). Finally, the frequency of SNP 282 C>T was significantly higher than in Spanish and European descents ( $p<0.05$ ) while that of SNP 590 G>A was significantly higher only when compared with Amerindians ( $p<0.001$ ).

#### 4. Discussion

In the present study, by sequencing the entire *NAT2* coding region in clinical samples collected from Brazilians resident in two different regions of the country, we



Table 4  
Allelic frequency of the most common SNPs of *NAT2* human gene in different populations

Population	SNP (%)						
	191A	282T	341C	481T	590A	803G	857A
Brazil (808 alleles)	4.1 <sup>R</sup>	36.8 <sup>R</sup>	37.7 <sup>R</sup>	35.4 <sup>R</sup>	26.4 <sup>R</sup>	40.3 <sup>R</sup>	4.5 <sup>R</sup>
European descents <sup>a</sup> (2068 alleles)	1 <sup>**</sup>	25.8 <sup>**</sup>	45.9	45.6 <sup>φ</sup>	26.4	47	1.2 <sup>**</sup>
German <sup>b</sup> (1688 alleles)	0.1 <sup>*</sup>	30.8	46.5 <sup>φ</sup>	42.5	27.8	42.4	1.3 <sup>φ</sup>
Polish <sup>c</sup> (496 alleles)	0 <sup>*</sup>	33.5	44.4	38.3	30	39.3	3.4
Spanish <sup>d</sup> (1008 alleles)	1.6	27.3 <sup>φ</sup>	45	44.8 <sup>φ</sup>	26	47.2	1.2 <sup>φ</sup>
Thai <sup>e</sup> (470 alleles)	ND	47.4	ND	3.8 <sup>*</sup>	32.5	ND	20.4 <sup>*</sup>
Japanese <sup>f</sup> (96 alleles)	0	37.3	1.6 <sup>*</sup>	1.6 <sup>*</sup>	28.3	7.4 <sup>*</sup>	11.3
Chinese <sup>g</sup> (882 alleles)	ND	30.5	6 <sup>*</sup>	6 <sup>*</sup>	30.5	ND	11.2 <sup>**</sup>
Amerindians							
Ngawbe <sup>h</sup> (210 alleles)	ND	25.2	2.4 <sup>*</sup>	2.4 <sup>*</sup>	0 <sup>*</sup>	1.9 <sup>*</sup>	23.3 <sup>*</sup>
Embera <sup>h</sup> (272 alleles)	ND	29	9.9 <sup>*</sup>	9.2 <sup>*</sup>	3.7 <sup>*</sup>	9.9 <sup>*</sup>	22.8 <sup>*</sup>
African <sup>i</sup> (234 alleles)	6.8	32.5	35.9	29.5	27.4	46.5	2.2
Sub-Saharan African <sup>j</sup> (250 alleles)	19.2 <sup>*</sup>	ND	ND	ND	ND	ND	ND
Effect on enzymatic activity	Decrease	No effect	Decrease	No effect	Decrease	No effect	Decrease

ND: not determined. Bonferroni correct *p*-value: R: reference. <sup>φ</sup>*p*-value 0.1–0.05. <sup>φ</sup>*p*-value 0.05–0.01. <sup>\*\*</sup>*p*-value 0.01–0.001. <sup>\*</sup>*p*-value < 0.001.

<sup>a</sup> [24].

<sup>b</sup> [21].

<sup>c</sup> [39].

<sup>d</sup> [23].

<sup>e</sup> [27].

<sup>f</sup> [26].

<sup>g</sup> [25].

<sup>h</sup> [28].

<sup>i</sup> [39] pooled data of Gabonese and Dogon populations.

<sup>j</sup> [40] Guinea Bissau population.

described the frequency of 13 SNPs, seven of which are those commonly found in high frequencies in ethnic closed populations in the world. Four of them (191 G > A, 341 T > C, 590 G > A, 857 G > A) are associated with the slow acetylation phenotype and three (282 C > T, 481 C > T and 803 A > G) have been demonstrated not to interfere with the acetylation status. Six SNPs (70 T > A; 345 C > T; 403 C > G; 609 G > T; 683 C > T and 838 G > A), that do not take part of the official *NAT2* nomenclature, were found in frequencies lower than 1% in the studied groups and their influence on the acetylation phenotype has not been established yet. However, except for the 345 C > T the remaining five are African-specific SNPs corroborating with the natural history of the African admixture in the constitution of the Brazilian population [36,37]. Analysis of our samples showed that although some SNPs (70 T > A, 345 C > T and 683 C > T) are not detected in Goiás, no statistically significant difference was observed when the frequencies of any the SNP identified were compared between the two regions. Additionally, two SNPs (282 C > T and 481 C > T, both silent mutations) were observed in high frequencies of 37% and 35%, respectively. The reasons why

these silent SNPs belong to the group of the most frequently distributed around the world is not well known. In our population, SNP 282 C > T was found in frequencies significantly higher than in European descents and Spanish (*p* < 0.01 and < 0.05, respectively). This may be interpreted as one of the consequences of the influence of Amerindians and Africans in the constitution of the Brazilian population.

Several SNPs with recognized role in the determination of the acetylation phenotype have been observed with significant high frequencies within certain ethnic groups. The variation 191 G > A (which characterizes the cluster *NAT2\*14*), is more prevalent in African populations whereas the SNP 857 G > A is predominant in Asians and Amerindians. The SNPs 341 T > C (cluster *NAT2\*5*), 481 C > T and 803 A > G are usually present in high frequencies among Europeans and Africans while SNP 590 G > A (cluster *NAT2\*6*) is more prevalent in Asians [21–28,38–43]. The finding of SNPs (191 G > A, 341 T > C and 857 G > A) of different ethnic-origin in two different regions of Brazil corroborates with the multiple ancestry of Brazilian population and clearly shows the influence of different ethnicities in the gene

pool of Brazilians, a trihybrid population with European, African and Amerindian roots.

Upon submitting the 13 SNPs identified in the present study to a linkage disequilibrium test (LDT), and after the exclusion of 4 SNPs that were not in LD with any other, 16 different alleles were characterized. Five of which have not been described yet, in consequence, no information is available on their functional activity. Among the remaining 11 alleles, the majority ( $n = 8$ ) were defective (associated with slow acetylation phenotype), from which *NAT2\*5B* and *\*6A* were the most frequent. The *NAT2\*4* was the most frequent one among the functional alleles and upon comparing frequencies of all alleles in individuals from Rio de Janeiro and Goiás, only the wild type allele showed a statistically significant difference being higher in Goiás. Additionally, samples from Goiás presented lower allele diversity than those from Rio de Janeiro. Despite of this and taking into account the overall allele distribution, there was a general predominance of defective alleles in both regions.

It was observed that the slow acetylation phenotype is mainly characterized by the *NAT2\*5B* allele (33%), followed by *NAT2\*6A* (26%), whereas the *NAT2\*4* was the most frequent among the functional alleles (20%). These findings once more suggest the influence of Europeans and Africans in the composition of the Brazilian's genetic background, considering the high incidence of *NAT2\*5* and *NAT2\*6* alleles among these two populations [22,42]. In Asians, both *NAT2\*6* and *NAT2\*7* alleles are the main responsible for the low acetylation phenotypes while the *NAT2\*5* cluster is rare [26,27]. Finally, in Amerindians, *NAT2\*7B* is the main allele responsible for the low acetylation status, with low incidence or absence of other low activity alleles [28]. The *NAT2\*7* was observed at low frequency (4%) among Brazilian individuals of both regions.

A total of 120 DNA samples were either homozygous for all SNPs or heterozygous for only one SNP, allowing the definition of the haplotype pairs. However, after the exclusion of six SNPs that were not informative for phenotype prediction, we were able to characterize 202 individuals' genotypes and predict their acetylation status. We obtained a total of 14 different genotypes with a predominance of the slow acetylation status (49%) represented mainly by *NAT2\*5/\*5* and *NAT2\*6/\*6*. Furthermore, the two regions studied here showed a different distribution of genotypes and acetylation status with a statistically significant higher frequency of the intermediate acetylation phenotype in Rio de Janeiro.

Further analysis of these two regions showed a lower diversity of allele and genotype variants in Goiás, probably as a consequence of the high influence of Amerindian

population (already characterized by a decreased genetic diversity and reduced genotype variability) in this region, according with the history of Brazil's colonization. The high migratory movement of different population groups along the Brazilian territory results in a considerable phylogeographical heterogeneity, probably influencing the distribution of allelic variants of *NAT2* among the different geographical regions of the country. Therefore, characterization of *NAT2* alleles profile in different regions of Brazil would be of high importance.

The high frequency of genotypes with slow acetylation activity alleles observed in the two studied regions of Brazil indicates a significant prevalence of the slow acetylation phenotypes, suggesting a higher probability of hepatic injury in TB patients upon INH treatment. The data we present on the genetic diversity among people from two different regions of Brazil also demonstrate the importance of genotyping *NAT2* in other regions since anti-TB medications in Brazil are standardized for the whole country. Prediction of hepatotoxicity phenotypes based on genotypes in different regions of Brazil may help the application of more appropriate therapeutic schemes to avoid adverse effects and to reduce treatment cost.

In conclusion, this first evaluation of *NAT2* gene in Brazilians subjects demonstrated a considerable prevalence of slow acetylators over rapid acetylators and established a basis for further detailed clinical and epidemiological investigations of *NAT2* and other conditions associated with the acetylation polymorphism in the different ethnic groups of the country.

## Acknowledgments

This work was supported by MCT/DECIT/MS and Faperj/Pronex Proc: E-26/171.207/2003. Raquel L.F. Teixeira was fellow awarded by FAPERJ/FIOCRUZ program. Adalberto R. Santos is supported by CNPq Grant number 308786/2005-0. We thank the nursing service of the Anuar Auad Infectious Disease Reference Hospital, Santa Maria Hospital from the State of Rio de Janeiro and from the Hospital Tuberculosis Control Program/Thoracic Diseases Institute/Clementino Fraga Filho University Hospital, for sample collection. Additionally we also thank the PDTIS-FIOCRUZ Sequencing and Bioinformatics Platforms support.

## References

- [1] M. Ingelman-Sundberg, Pharmacogenetics: an opportunity for a safer and more efficient pharmacotherapy, *J. Intern Med.* 250 (2001) 186–200.

- [2] J. Lundkvist, B. Jonsson, Pharmacoeconomics of adverse drug reactions, *Fundam. Clin. Pharmacol.* 18 (2004) 275–280.
- [3] A. Marshall, Getting the right drug into the right patient, *Nat. Biotechnol.* 15 (1997) 1249–1252.
- [4] R.A. Garibaldi, R.E. Drusin, S.H. Ferebee, M.B. Gregg, Isoniazid-associated hepatitis. Report of an outbreak, *Am. Rev. Respir. Dis.* 106 (1972) 357–365.
- [5] Y.S. Huang, H.D. Chern, W.J. Su, J.C. Wu, S.L. Lai, S.Y. Yang, F.Y. Chang, S.D. Lee, Polymorphism of the *N*-acetyltransferase 2 gene as a susceptibility risk factor antituberculosis drug-induced hepatitis, *Hepatology* 35 (2002) 883–889.
- [6] D.W. Hein, M.A. Doll, A.J. Fletland, M.A. Leff, S.J. Webb, G.H. Xiao, U.S. Devanaboyina, N.A. Nangju, Y. Feng, Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms, *Cancer Epidemiol. Biomarkers Prev.* 9 (2000) 29–42.
- [7] D.W. Hein, Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis, *Mutat. Res.* 506–507 (2002) 65–77.
- [8] S. Boukouvala, E. Sim, Structural analysis of the genes for human arylamine *N*-acetyltransferases and characterization of alternative transcripts, *Basic Clin. Pharmacol. Toxicol.* 96 (2005) 343–351.
- [9] H.B. Hughes, J.P. Biehl, A.P. Jones, L.H. Schmidt, Metabolism of isoniazid in man as related to the occurrence of peripheral neuritis, *Am. Rev. Respir. Dis.* 70 (1954) 266–273.
- [10] D.A.P. Evans, T.A. White, Human acetylation polymorphism, *J. Lab. Clin. Med.* 63 (1964) 394–403.
- [11] W.W. Weber, D.W. Hein, Clinical pharmacokinetics of isoniazid, *Clin. Pharmacokinet.* 4 (1979) 401–422.
- [12] I. Cascorbi, J. Brockmüller, P.M. Mrozikiewicz, S. Bauer, R. Loddenkemper, I. Roots, Homozygous rapid arylamine *N*-acetyltransferase (NAT2) genotype as a susceptibility factor for lung Cancer, *Cancer Res.* 56 (1996) 3961–3966.
- [13] A.C. Deitz, W. Zheng, M.A. Leff, M. Gross, W.Q. Wen, M.A. Doll, G.H. Xiao, A.R. Folsom, D.W. Hein, *N*-acetyltransferase-2 genetic polymorphism, well-done meat intake, and breast cancer risk among postmenopausal women, *Cancer Epidemiol. Biomarkers Prev.* 9 (2000) 905–910.
- [14] M. Ohno, I. Yamaguchi, I. Yamamoto, T. Fukuda, S. Yolota, R. Maekura, M. Ito, Y. Yamamoto, T. Ogura, K. Maeda, K. Komuta, T. Igarashi, J. Azuma, Slow *N*-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity, *Int. J. Tuberc. Lung Dis.* 4 (2000) 256–261.
- [15] T. Kita, Y. Tanigawara, S. Chikazawa, H. Hatanaka, T. Sakaeda, F. Komada, S. Iwakawa, K. Okumura, *N*-acetyltransferase 2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients, *Biol. Pharm. Bull.* 24 (2001) 544–549.
- [16] Y.S. Huang, H.D. Chern, J.C. Wu, Y. Chao, Y.H. Huang, F.Y. Chang, S.D. Lee, Polymorphism of the *N*-acetyltransferase 2 gene, red meat intake, and the susceptibility of hepatocellular carcinoma, *Am. J. Gastroenterol.* 98 (2003) 1417–1422.
- [17] H.W. Harris, R.A. Knight, M.L. Selin, Comparison of isoniazid concentrations in the blood of people of Japanese and European descent; therapeutic and genetic implications, *Am. Rev. Tuberc.* 78 (1958) 944–948.
- [18] D.A.P. Evans, *N*-acetyltransferase, *Pharmacol. Ther.* 42 (1989) 157–234.
- [19] A.J. Fretland, M.A. Leff, M.A. Doll, D.W. Hein, Functional characterization of human *N*-acetyltransferase 2 (NAT2) single nucleotide polymorphisms, *Pharmacogenetics* 11 (2001) 207–215.
- [20] H.J. Lin, C.H. Han, B.K. Lin, S. Hardy, 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.* 52 (1993) 827–834.
- [21] I. Cascorbi, N. Drakoulis, J. Brockmüller, A. Maurer, K. Sperling, I. Roots, Arylamine *N*-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity, *Am. J. Hum. Genet.* 57 (1995) 581–592.
- [22] C. Dandara, C.M. Masimirembwa, A. Magimba, S. Kaaya, J. Sayi, D.K. Sommers, J.R. Snyman, J.A. Hasler, Arylamine *N*-acetyltransferase (NAT2) genotypes in Africans: the identification of a new allele with nucleotide changes 481 C > T and 590 G > A, *Pharmacogenetics* 13 (2003) 55–58.
- [23] J.A.G. Agúndez, M. Oliveira, C. Martinez, J.M. Ladero, J. Benítez, Identification and prevalence study of 17 allelic variants of human NAT2 gene in a White population, *Pharmacogenetics* 6 (1996) 423–428.
- [24] J.A.G. Agúndez, NAT2 Genotyping: Equilibrium between accuracy and feasibility in routine analyses, *J. Appl. Res.* 3 (2003) 1–6.
- [25] H.G. Xie, Z.H. Xu, D.S. Ou-Yang, Y. Shu, D.L. Yang, J.S. Wang, X.D. Yan, S.L. Huang, W. Wang, H.H. Zhou, Meta-analysis of phenotype and genotype of NAT2 deficiency in Chinese populations, *Pharmacogenetics* 7 (1997) 503–514.
- [26] A. Sekine, S. Saito, A. Iida, Y. Mitsunobu, S. Higuchi, S. Harigae, Y. Nakamura, Identification of single-nucleotide polymorphisms (SNPs) of human *N*-acetyltransferase genes NAT1, NAT2, AANAT, ARD1 and L1CAM in the Japanese population, *J. Hum. Genet.* 46 (2001) 314–319.
- [27] V. Kukongviriyapan, A. Prawan, W. Tassaneyakul, J. Aiemsard, B. Warasiha, Arylamine *N*-acetyltransferase-2 genotypes in the Thai population, *Br. J. Clin. Pharmacol.* 55 (2003) 278–281.
- [28] L.F. Jorge-Nebert, M. Eichelbaum, E.U. Griese, T. Inaba, T.D. Arias, Analysis of six SNPs of NAT2 in Ngawbe and Embera Amerindians of Panama and determination of the Embera acetylation phenotype using caffeine, *Pharmacogenetics* 12 (2002) 39–48.
- [29] C. Martinez, J.A.G. Agundez, M. Olivera, A. Llerena, R. Ramirez, M. Hernandez, J. Benitez, Influence of genetic admixture on polymorphisms of drug-metabolizing enzymes: analyses of mutations on NAT2 and CYP2E1 genes in a mixed Hispanic population, *Clin. Pharmacol. Ther.* 63 (1998) 623–628.
- [30] F.C. Parra, R.C. Amado, J.R. Lambertucci, J. Rocha, C.M. Antunes, S.D.J. Pena, Color and genomic ancestry in Brazilians, *Proc. Natl. Acad. Sci. U.S.A.* 100 (2003) 177–182.
- [31] S. Fuselli, R.H. Gilman, S.J. Chanock, S.L. Bonatto, G. De Stefano, C.A. Evans, D. Labuda, D. Luiselli, F.M. Salzano, G. Soto, G. Vallejo, A. Sajantila, D. Pettener, E. Tarazona-Santos, Analysis of nucleotide diversity of NAT2 coding region reveals homogeneity across Native American populations and high intra-population diversity, *Pharmacogenomics J.* (2006) 1–9 [Epub ahead of print].
- [32] G. Warnes, F. Leisch, Genetics: Population Genetics. R package version 1.2.0. (2005).
- [33] J.P. Sinnwell, D.J. Schaid, Haplo.stats: Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. R package version 1.2.2 (2005), <http://mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm>.
- [34] D.J. Schaid, C.M. Rowland, D.E. Tines, R.M. Jacobson, G.A. Poland, Score tests for association between traits and haplotypes when linkage phase is ambiguous, *Am. J. Hum. Genet.* 70 (2002) 425–434.

- [35] R Development Core Team. R: A language and environment for statistical computing. (2005) R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.r-project.org>.
- [36] E. Patin, C. Harmant, K.K. Kidd, J. Kidd, A. Froment, S.Q. Mehdi, L. Sica, E. Heyer, L. Quintana-Murci, Sub-Saharan African coding sequence variation and haplotype diversity at the *NAT2* gene, *Hum Mutat.* 27 (2006) 720.
- [37] E. Patin, L.B. Barreiro, P.C. Sabeti, F. Austerlitz, F. Luca, A. Sajantila, D.M. Behar, O. Semino, A. Sakuntabhai, N. Guiso, B. Gicquel, K. McElreavey, R.M. Harding, E. Heyer, L. Quintana-Murci, Deciphering the ancient and complex evolutionary history of human arylamine *N*-acetyltransferase genes, *Am. J. Hum. Genet.* 78 (2006) 423–436.
- [38] D.A. Bell, J.A. Taylor, M.A. Buttler, E.A. Stephens, J. Wiest, L.H. Brubaker, F.F. Kadlubar, G.W. Lucier, Genotype/phenotype discordance for human arylamine *N*-acetyltransferase (*NAT2*) reveals a new slow-acetylator allele common in African-Americans, *Carcinogenesis* 14 (1993) 1689–1692.
- [39] C. Delomenie, L. Sica, D.M. Grant, R. Krishnamoorthy, J.M. Dupret, Genotyping of the polymorphic *N*-acetyltransferase (*NAT2\**) gene locus in two native African populations, *Pharmacogenetics* 6 (1996) 177–185.
- [40] I. Cavaco, R. Reis, J.P. Gil, V. Ribeiro, *CYP3A4\*1B* and *NAT2\*14* alleles in a native African population, *Clin. Chem. Lab. Med.* 41 (2003) 606–609.
- [41] M. Blum, A. Demierre, D.M. Grant, M. Heim, U.A. Meyer, Molecular mechanism of slow acetylation of drugs and carcinogens in human, *Proc. Natl. Acad. Sci. U.S.A.* 88 (1991) 5237–5241.
- [42] P.M. Mrozikiewicz, I. Cascorbi, J. Brockmöller, I. Roots, Determination and allelic allocation of seven nucleotide transitions within the arylamine *N*-acetyltransferase gene in the Polish population, *Clin. Pharmacol. Ther.* 59 (1996) 376–382.
- [43] A.S. Aynacioglu, I. Cascorbi, P.M. Mrozikiewicz, I. Roots, Arylamine *N*-acetyltransferase (*NAT2*) genotypes in a Turkish population, *Pharmacogenetics* 7 (1997) 327–331.



## CAPÍTULO II

Essa etapa incluiu a redação de um pedido de patente nacional, relacionada ao campo da Biologia Molecular e Genômica, especialmente a Farmacogenômica, que inclui metodologia passível de aplicação para orientação de condutas terapêuticas utilizando os novos polimorfismos descritos.

A referida patente intitula-se “Método para a predição do desfecho do tratamento de doenças humanas utilizando fármacos metabolizados pela N-acetiltransferase 2 humana (NAT2) com base em polimorfismos genéticos” sendo depositada no INPI (Instituto Nacional de Propriedade Industrial) na data de 04-10-07 onde recebeu seu número de depósito BRPI0704093-8. Os inventores são: Raquel Lima de F. Teixeira, Adalberto Rezende Santos, Antonio Basílio de Miranda e Philip Noel Suffys da FIOCRUZ e Fernanda C. Q. Mello e Afrânio Lineu Kritski da UFRJ.

Em outubro de 2008, segundo informação da GESTEC, a Fiocruz estendeu o pedido de patente nacional para patente internacional, a ser protegida em alguns mercados ainda por serem definidos. O número do pedido de patente depositado no exterior via sistema PCT (*Patent Cooperation Treaty*), é PCT/BR2008/000304, depositado em 06-10-2008.

## RESUMO

**"Método para a predição do desfecho do tratamento de doenças humanas utilizando fármacos metabolizados pela N-acetiltransferase 2 humana (NAT2) com base em polimorfismos genéticos"**

A presente invenção está relacionada ao campo da Biologia Molecular e Genômica, especialmente a Farmacogenômica. A invenção descreve a presença de novos polimorfismos no gene que codifica para a enzima Arilamina N-acetiltransferase 2 humana (NAT2), a qual é responsável pela metabolização de fármacos importantes na terapêutica de várias doenças de etiologias diversas, bem como de inúmeras toxinas e carcinógenos presentes em alimentos, cigarro e no ambiente. A presente invenção inclui metodologia passível de utilização na terapêutica utilizando os polimorfismos descritos.

“Método para a predição do desfecho do tratamento de doenças humanas utilizando fármacos metabolizados pela N-acetiltransferase 2 humana (NAT2) com base em polimorfismos genéticos”

## 5 **Campo da Invenção**

A presente invenção está relacionada ao campo da Biologia Molecular e Genômica, especialmente a Farmacogenômica. Descreve a presença de novos polimorfismos no gene que codifica para a enzima Arilamina N-acetiltransferase 2 humana (NAT2), a qual é responsável pela metabolização de fármacos importantes na terapêutica de várias doenças de etiologias diversas, bem como de inúmeras toxinas e carcinógenos presentes em alimentos, cigarro e no ambiente. A presente invenção inclui metodologia passível de utilização na terapêutica utilizando os polimorfismos descritos.

## **Fundamentos e Estado da Arte**

A análise de variações na seqüência de DNA para a compreensão dos diferentes fenótipos observados no ser humano tem sido o principal foco da pesquisa genética durante décadas. Diferentes testes têm sido extensivamente explorados através da análise de ligações, segundo os quais, mutações em um único gene podem ser suficientes não somente para explicar variações na suscetibilidade de diferentes indivíduos a uma ou mais doenças como também para explicar variações inter-individuais de resposta a um determinado fármaco. O tipo mais comum de variação encontrada no genoma humano é o polimorfismo de base única (*Single Nucleotide Polymorphism* - SNP). Tais variações têm sido identificadas como marcadores genéticos

para doenças de etiologias diversas sendo também relacionados com diferentes fenótipos de metabolização de xenobióticos.

Ingelman-Sundberg, J Int Med 250: 186-200 2001; Roses, A.D Life Sciences 70:1471-1480 2002 e Prows & Prows, American  
5 Journal of Nursing 104:60-70 2004, através de estudos farmacogenéticos, nos quais foram avaliadas variações genéticas que causam diferentes respostas aos medicamentos, ao incluírem dados acerca das diferenças na eficácia, interações fármaco-fármaco e o risco relativo de reações adversas,  
10 identificaram uma associação de vários SNPs com a expressão ou atividade de receptores farmacológicos, proteínas transportadoras, sinalizadoras e enzimas do metabolismo.

Um dos principais objetos de estudo da farmacogenética refere-se a análise de enzimas envolvidas na biotransformação  
15 de fármacos. Variações nos genes que codificam para essas enzimas podem resultar em uma diminuição de atividade destes fármacos, e favorecer o aparecimento de reações adversas por toxicidade devido a uma baixa metabolização e conseqüente acúmulo do princípio ativo ou de metabólitos intermediários  
20 tóxicos (metabolizadores lentos). Por outro lado, segundo Ingelman-Sundberg, J Int Med 250: 186-200 2001, a presença de variantes alélicas que não afetam a atividade enzimática, isto é, alelos funcionais (que caracterizam um fenótipo de metabolizadores rápidos) ou com múltiplas cópias  
25 (metabolizadores ultra-rápidos) pode estar associada a concentrações séricas subterapêuticas do fármaco. Adicionalmente, de acordo com Kalow, Pharmacological Reviews 49:369-379 1997, nos casos de heterozigose onde os indivíduos sejam portadores de um alelo que caracterize baixa atividade



enzimática e um alelo funcional, um fenótipo de acetilação intermediária é originado.

A enzima N-acetiltransferase 2 humana (NAT2), codificada pelo gene *NAT2* está envolvida na metabolização de vários fármacos utilizados correntemente para o tratamento de: 5 tuberculose (isoniazida), hipertensão arterial (hidrazinas anti-hipertensivas), hanseníase (dapsona) e arritmia cardíaca (procainamida anti-arritmica) além de sulfametazina e outras sulfonamidas usadas para a terapia de outras desordens (Weber, 10 *The Acetylator Genes and Drug Response*, Oxford University Press, N.Y. 1987). Várias mutações pontuais na região codificante do gene *NAT2* são capazes de alterar a atividade de acetilação da enzima podendo resultar em três diferentes fenótipos: acetiladores lentos, intermediários e rápidos 15 (Hein, et al., 2000. *Cancer Epidemiology, Biomarkers and Prevention*, 9:29-42). O fenótipo de acetilação lenta, caracterizado por determinados perfis genotípicos leva a um acúmulo do fármaco e conseqüentemente ao desenvolvimento de reações adversas, as quais são principalmente caracterizadas 20 por neuropatias periféricas e hepatite. Já o fenótipo de acetilação rápida leva a uma diminuição dos níveis séricos do fármaco, as quais podem baixar até alcançar níveis sub-terapêuticos levando conseqüente à uma falência terapêutica (Ingelman-Sundberg, *J Int Med* 250: 186-200 2001).

25 A enzima NAT2 participa também de vias de ativação de poluentes ambientais com potencial carcinogênico, tais como, 2-aminofluoreno, 4-aminobifenil, benzidina, beta-nafitilamina e algumas aminas heterocíclicas presentes em pirolisados protéicos (Kato, *CRC Crit Ver.Toxicol.* 16:307-348 1986; Weber,

The Acetylator Genes and Drug Response, Oxford University Press, N.Y. 1987; Hein *Biochim Biophys Acta*, 948:37-66, 1988). Adicionalmente, a participação de NAT2 na detoxificação também tem sido associada à suscetibilidade a doenças induzidas por substâncias químicas, tais como neoplasia (Vatsis et al, *Pharmacogenetics*, 5: 1-17 1995) quando da combinação da ingestão de carne vermelha com o fumo (Potter et al *Câncer Epidem. Biomarkers & Prev.*, 8: 69-75 1999; Liu et al., *Canc Letters*, 133: 115-123 1988).

10 Até o presente momento, já foram identificados 36 alelos diferentes para o gene *NAT2* (<http://www.louisville.edu/medschool/pharmacology/NAT.html>), os quais consistem na combinação de até 4 mutações pontuais na região codificante do gene. Estudos funcionais realizados no início desta década demonstraram a existência de diferentes mecanismos que podem resultar no fenótipo de acetilação lenta dependendo dos polimorfismos presentes no gene *NAT2*. Dentre estes mecanismos estão: (i) diminuição da quantidade de proteína sintetizada, (SNPs C190T, T341C, A434C, G590A); (ii) diminuição da atividade enzimática (SNPs C190T, G191A, T341C, A434C e G59A) e (iii) diminuição da estabilidade da enzima *NAT2* (SNPs G191A, A845C e G857A), (Fretland et al., *Pharmacogenetics*, 11:207-215 2001; Zhu & Hein, *Biol. Chem.*, 383: 983-987 2002). Por outro lado, outras mutações descritas na literatura como por exemplo T111C, C282T, C481T, C759T e A803G, não exercem nenhum efeito na atividade biológica da proteína, caracterizando um alelo funcional (Fretland et al., *Pharmacogenetics*, 11:207-215 2001). Estudos adicionais que avaliaram a relação entre diferentes haplótipos e o desfecho

de metabolização, baseados na expressão e atividade enzimática realizados com os alelos mais frequentes de *NAT2*, possibilitaram a caracterização dos mesmos quanto ao perfil de acetilação. Os alelos dos clusters *NAT2* \*5, \*6, \*7, \*14 e \*19 conferem um fenótipo de acetilação lenta enquanto os alelos *NAT2* \*4 e dos clusters \*12 e \*13, conferem um fenótipo de acetilação rápida (Hein *et al*, *Cancer Epidemiology, biomarkers and Prevention*, 9:29-42 2000; Hein, *Mutation Research*, 506-507:65-77. 2002).

10 As Reações Adversas a Drogas (ADRs) são causas comuns de hospitalização em todo o mundo e passíveis de promover o óbito de pacientes acometidos por tais situações. Recentemente, Bagheri *et al*, *J. Clin. Pharmacol.* 50: 479-484 2000; Bissell *et al*, *Hepatology* 33:1009-1013 2001; Larrey, *Seminars in Liver Disease* 22:145-155 2002, descreveram que, entre as reações adversas, a lesão hepática é o tipo mais frequente. Segundo os autores, a doença hepática induzida por fármacos é uma das principais complicações durante o uso de antimicrobianos, uma vez que o fígado é o órgão central na biotransformação e excreção da maioria dos fármacos e substâncias xenobióticas.

20 Variações genéticas em enzimas que participam na geração destes bioprodutos tóxicos podem resultar em diferenças na reatividade metabólica a agentes específicos com variações de indivíduo para indivíduo. Assim, a determinação dos perfis genotípicos individuais e seus fenótipos resultantes podem ser de grande valia para os médicos tanto na escolha do fármaco a ser indicado quanto na determinação da posologia e do esquema terapêutico a ser adotado, o qual deve encontrar um balanço adequado entre a eficácia e a toxicidade. Adicionalmente,

estas variações genéticas podem ser utilizadas como marcadores para identificar pré-disposição a cânceres.

Prince & Brookes, Expert Rev. Mol. Diagn. 1(3) 89-95,2001. "Towards high-throughput genotyping of SNPs by dynamic allele-specific hybridization", descreveram a análise de SNPs como  
5 uma estratégia simples para o descobrimento de genes envolvidos em doenças complexas tais como Alzheimer, obesidade e diabetes melito e anteciparam que os SNPs teriam um papel muito importante na farmacogenômica, onde a identificação de  
10 variações em genes específicos, relevantes na eficácia de determinada droga, toxicidade e metabolismo ajudariam a estabelecer estratégias terapêuticas ótimas para pacientes individuais. Refletindo estas expectativas, várias novas tecnologias têm aparecido ao longo dos últimos anos, cada uma  
15 com vantagens particulares, mas todas com o objetivo único de simplificar a análise dos SNPs. Os autores descrevem uma técnica chamada "Hibridação dinâmica alelo-específica" (DASH), segundo eles, um método conveniente para a genotipagem de SNPs e mutações do tipo *indel*, (inserção ou deleção), as quais  
20 possuem potencial de aplicação tanto na pesquisa básica como no diagnóstico clínico. Informam ainda que o DASH comercial está disponível, tornando a tecnologia acessível para todos os laboratórios.

Garte, Cancer Epidemiology Biomarkers & Prevention Vol.10,  
25 1233-1237, December 2001. "Metabolic Susceptibility Genes As Cancer Risk Factors" descrevem que polimorfismos em genes da via metabólica relacionada à ativação de carcinógenos que conferem suscetibilidade ao câncer não mostraram uma associação consistente com o desenvolvimento de câncer quando

somente os efeitos principais foram examinados. Isto realmente era esperado tendo em vista o limitado e específico papel bioquímico de tais genes no processo carcinogênico. Contudo, quando grupos particulares de casos são examinados separadamente, a importância destes polimorfismos genéticos torna-se clara. Exemplos da literatura e um modelo hipotético são apresentados como suporte de que alelos de risco pertencentes à via metabólica devem ser estudados em subgrupos de estudos do tipo caso controle com maior casuística tendo como hipótese questões relacionadas à ação do produto gênico em função de variáveis demográficas, ambientais ou outras variáveis genéticas.

O pedido de patente US 20030096235 "Methods and products related to genotyping and DNA analysis", descreve métodos e produtos relacionados a genotipagem. Este método é baseado no uso de polimorfismos de base única (SNPs) para uma varredura ampla e de alta *performance* do genoma. O método é realizado através da hibridização de oligonucleotídeos alelo específicos em um genoma de complexidade reduzida (RCG). A invenção também se relaciona com métodos de: preparo de oligonucleotídeos SNP-específicos e RCGs, métodos de *fingerprinting*, determinação da frequência alélica para um determinado SNP, caracterização de tumores, geração de um código de classificação genômico para a identificação de SNPs desconhecidas, composições relacionadas e *kits*.

O pedido de patente US 20030049628 "Complexity Management of Genomic DNA", descreve métodos e *kits* para reduzir a complexidade de uma amostra de DNA através da disponibilização de métodos (não baseados em gel) para a amplificação de um

subgrupo de seqüências em uma amostra com sítios de restrição para duas ou mais enzimas, através da ligação de adaptadores aos fragmentos, de forma que somente um subgrupo de fragmentos possa ser amplificado. Esta invenção descreve adicionalmente  
5 procede a análise do produto amplificado acima por hibridação com um arranjo que pode ser especificamente desenhado para características particulares dos fragmentos desejados, tais como, a presença ou ausência de um polimorfismo.

O pedido de patente US 20040053232 "Quantitative trait  
10 loci and somatostatin", relaciona-se ao uso de características genéticas em gado para a determinação das características da progênie e para otimizar a manipulação do gado visando um aumento na performance de amamentação e qualidade da carne. O alvo especialmente estudado é o *locus* da somatostatina, o qual  
15 está associado com certos *loci* de características quantitativas (QTLs). Os SNPs e haplótipos são preditivos do aumento ou decréscimo da quantidade de tecido no animal.

O pedido de patente US 20040081996 "Methods of validating  
20 SNPs and compiling libraries of assays", descreve bibliotecas de ensaios e métodos para a compilação de dados em bibliotecas. Os ensaios podem identificar polimorfismos de base única (SNPs). Adicionalmente, métodos de validação de SNPs e construção de mapas de desequilíbrio de ligação usando conjuntos ou sub-conjuntos de SNPs são disponibilizados.

25 O pedido de patente US 20050042654 "Genotyping Methods" descreve métodos para a amplificação do DNA genômico e genotipagem das amostras amplificadas. Os métodos de genotipagem utilizam arranjos de oligonucleotídeos alelo específicos para polimorfismos de base única (SNPs). Os

métodos também relacionam técnicas para a amplificação de múltiplas amostras de DNA de vários indivíduos de maneira a minimizar a possibilidade de contaminação das amostras por produtos amplificados previamente (amplicons).

5 O pedido de patente US 20050153328 "Method and markers for determining the genotype of horned/polled cattle", descreve métodos para a descoberta e utilização de polimorfismos de base única (SNPs) na determinação do genótipo de ruminantes. Adicionalmente, a presente invenção fornece seqüências  
10 específicas de ácidos nucleicos, SNPs e perfis de SNPs que podem ser usados com esta finalidade.

Chang-Claude et al., Cancer Epidemiology, Biomarkers & Prevention, 11 698-704, 2002. "Differential Effect of NAT2 on the Association between Active and Passive Smoke Exposure and  
15 Breast Cancer Risk", em estudo caso-controle realizado para avaliar o possível efeito diferencial do genótipo de NAT2 na relação entre fumantes ativos e passivos e o câncer de mama, descreveram que polimorfismos presentes no gene que codifica para a enzima N-acetyltransferase 2 (NAT2) influenciam na taxa  
20 de metabolização de amins aromáticas heterocíclicas presentes no tabaco.

Anitha & Banerjee, Int J Mol Med Vol.11 125-131, 2003. "Arylamine N-acetyltransferase 2 polymorphism in the ethnic populations of South India ", mostraram que o polimorfismo da  
25 N-acetilação é um trato genético refletido fenotipicamente pelas diferenças na atividade da N-acetyltransferase 2 (NAT2) em agentes terapêuticos (acetiladores rápidos e lentos). Os autores relatam os resultados da genotipagem de NAT2 de comunidades etnicamente diferentes do sul da Índia. Após a

genotipagem de 8 diferentes comunidades étnicas, o fenótipo de acetilação lenta teve predominância de 74% nestas populações corroborando resultados encontrados em outras populações asiáticas onde esta predominância é de cerca de 60%. Os  
5 autores não encontraram associação entre o sexo e o fenótipo de acetilação.

Musbah et al, Squ Journal For Scientific Research: Medical Sciences 2003 Vol 5, No. 1-2, 9-14. "Distribution of arylamine N-acetyltransferase 2 (NAT2) genotypes among Omanis".descreve  
10 sistemas para a identificação da maioria dos polimorfismos conhecidos de NAT2, tais como: G191A, C282T, C341T, C481T, G590A, A803G e G857A usando a técnica de PCR-RFLP (Polymerase Chain Reaction-Restriction Fragment Length.Polymorphism).

Cavaco et al, Clin Chemical Lab Med ol.41(4) 606-609 2003  
15 "CYP3A4\*1B and NAT2 \*14 Allele in a Native African Population", relatam polimorfismos de base única nos genes que codificam para o citocromo P450 3A4 (CYP3A4) e para a N-acetiltransferase 2 (NAT2), principais mediadores do metabolismo de uma grande variedade de drogas de uso  
20 terapêutico, bem como de xenobióticos.

Lash et al, Breast Cancer Research Vol: 7 N°3 R.385-393 2005. "A case-only analysis of the interaction between N-acetyltransferase haplotypes and tobacco smoke in breast cancer etiology" descreve haplótipos de N-acetyltransferase 2  
25 envolvidos com câncer de mama. Mulheres que possuem alelos mutantes com um genótipo homozigoto possuem uma taxa de ativação metabólica mais baixa das arilaminas aromáticas, um dos constituintes do tabaco, identificado como carcinógeno. Os autores levantam a hipótese de que mulheres com câncer de mama



com o fenótipo de acetilação lenta possuem um risco aumentado para o desenvolvimento deste tipo de carcinoma associado com o fumo ativo ou exposição passiva.

O pedido de patente US 20020128215. "Novel sequence  
5 variants of the human N-acetyltransferase 2 (NAT2) gene and  
use thereof", descreve novos polimorfismos no gene *NAT2*, os  
quais podem estar envolvidos no metabolismo de drogas  
utilizadas para o tratamento de várias desordens bem como  
podem ser utilizados para a determinação de pré-disposição ao  
10 desenvolvimento de vários cânceres.

Em resumo, nenhum dos estudos já realizados e documentos  
de patentes encontrados revelaram métodos similares, baseados  
nos SNPs aqui descritos, para predizer os efeitos  
hepatotóxicos provenientes da metabolização de drogas no  
15 organismo de um ser humano, através de variações na seqüência  
do gene codificador para a enzima Arilamina N-  
acetiltransferase 2 (NAT2). Estes e outros aspectos vantajosos  
formam coletivamente os objetos da presente invenção, que  
serão descritos em mais detalhes a seguir.

## 20 **Sumário da Invenção**

A presente invenção inclui seqüências de ácidos nucleicos  
mostradas nas FIGS. 1A-1B e Tabela 2, com destaque aos sítios  
polimórficos no gene que codifica para *NAT2*. A presente  
invenção relaciona-se ainda com oligonucleotídeos, mostrados  
25 na tabela 1, que hibridam especificamente com diferentes  
regiões da seqüência mostrada nas FIGS. 1A-1B. Adicionalmente,  
a presente invenção descreve como os polimorfismos estão  
representados na população de duas regiões geográficas  
distintas do Brasil.

Métodos diagnósticos que utilizem toda ou parte das seqüências contendo os SNPs identificados nesta invenção. Por exemplo, tais ácidos nucléicos ou especificamente os SNPs, podem ser utilizados como métodos para a identificação de polimorfismos no gene de NAT2 que estejam associados à predisposição a várias doenças, incluindo, mais não limitadas a câncer de bexiga e câncer de colo e câncer de próstata.

Uma outra aplicação desta invenção inclui a possibilidade de criação de um protocolo de prognóstico para pacientes recebendo composições terapêuticas metabolizadas por NAT2, tais como: isoniazida, fenilzina, hidrazina, dapsona, procainamida, sulfametazina e outras sulfonamidas. O método inclui: a) identificação de pacientes recebendo uma destas drogas; b) determinação do fenótipo de acetilação (rápidos ou lentos) com base no genótipo; e c) conversão dos dados obtidos no passo (b) em um protocolo de prognóstico. O protocolo de prognóstico pode incluir a predição da eficácia da droga, a predição do prognóstico do paciente, a predição de interações entre drogas, e a predição de efeitos adversos.

A invenção também está relacionada com a identificação de diferenças entre indivíduos na metabolização de compostos exógenos, incluindo, mas não limitados a carcinógenos ou agentes mutagênicos, incluindo 2-aminofluoreno, 4-aminobifenil, benzidina, beta-nafitilamina, e certas arilaminas heterocíclicas presentes em pirolisados de proteínas.

Em outro contexto, a invenção descreve ainda a frequência dos polimorfismos de NAT2 em duas populações de regiões geográficas distintas do Brasil, permitindo, com base nestas

informações, tanto a identificação de grupos étnicos mais suscetíveis a doenças ou desordens diversas, como também subsidiar os profissionais de saúde na elaboração e/ou redefinição de condutas terapêuticas em especial na escolha do fármaco e sua dosagem a serem utilizados por um determinado grupo étnico.

Uma outra aplicabilidade desta invenção é a criação de um método para assistir no desenvolvimento de novos compostos terapêuticos através de ensaios clínicos. O método inclui: a) administração de um composto terapêutico a um indivíduo e avaliação de sua eficácia; b) determinação, através do genótipo incluindo os SNPs aqui descritos, se o indivíduo é um acetilador rápido ou lento; e c) determinar a partir dos passos (a) e (b), qual composição terapêutica será a mais efetiva para aquele genótipo em particular e qual estará mais associada aos efeitos adversos.

### **Descrição das Figuras**

As Figuras 1A-1B mostram a seqüência do gene *NAT2* (alelo selvagem), (SEQ ID NO:1). Estas figuras contém a seqüência nucleotídica do tipo selvagem e a seqüência de aminoácidos (SEQ ID NO:2) O nucleotídeo de número 1 refere-se a adenosina do códon de iniciação "ATG", o qual encontra-se sublinhado. A posição das bases referente aos seis novos SNPs descobertos na região codificante do gene encontram-se sublinhadas e em negrito e correspondem as substituições de bases listadas na Tabela 2. O códon de terminação está indicado por um asterisco (\*). Adicionalmente, as trocas de aminoácidos encontram-se sublinhadas e em negrito.

A seguir são apresentadas definições importantes para a perfeita compreensão do escopo da presente invenção.

**"Amplificação de ácidos nucleicos"** refere-se a métodos tais como a reação da polimerase em cadeia (PCR) e reação em cadeia da ligase (LCR). Estes métodos são bem conhecidos, amplamente disseminados e tanto os equipamentos como os reagentes para a execução da técnica de PCR estão comercialmente disponíveis. Os oligonucleotídeos utilizados para a amplificação de uma região específica do cromossomo são preferencialmente complementares e hibridam especificamente a seqüências na região cromossômica ou em regiões que flanqueiam a região alvo. As seqüências geradas por amplificação podem ser seqüenciadas diretamente ou ainda serem clonadas.

**"Sequenciamento de DNA"** refere-se a uma técnica utilizada para determinação exata da seqüência nucleotídica de um fragmento particular de DNA ou de um genoma inteiro. A ordem dos pares de bases na molécula de DNA fornece uma informação específica da seqüência de aminoácidos de uma proteína a ser produzida por um determinado gene.

**"Condições de estringência"** para hibridações referem-se as condições de temperatura e concentração de tampão que permite a hibridação de um ácido nucleico particular a outro ácido nucleico sendo que o primeiro ácido nucleico pode ser perfeitamente complementar ao segundo, ou pode somente compartilhar algum grau de complementariedade, menor que o perfeito. Por exemplo, algumas condições de alta estringência podem ser utilizadas, as quais distinguem perfeitamente ácidos nucleicos complementares daqueles com complementariedade menor.

**"Clonagem"** refere-se ao uso de técnicas de recombinação *in vitro* a fim de inserir um gene particular ou outra seqüência de DNA numa molécula que servirá como vetor. Para se obter sucesso na clonagem de um determinado gene, é necessário o uso de métodos para a geração de fragmentos de DNA, que sejam capazes de se ligar a moléculas vetores, com conseqüente introdução da molécula híbrida em uma célula hospedeira na qual poderá se replicar, para então proceder-se a seleção dos clones contendo o gene alvo.

10 **"Veículo de clonagem"** refere-se a um DNA de plasmídeo ou bacteriófago, ou outra seqüência de DNA, a qual seja capaz de se replicar numa célula hospedeira. O veículo (ou vetor de clonagem), é caracterizado pela presença de um ou mais sítios de restrição no qual cada seqüência de DNA pode ser cortada sem a perda da função biológica do DNA o qual pode conter um marcador útil para a identificação e seleção das células transformadas.

**"Gene"** refere-se a uma seqüência de DNA portando informação genética. Esta pode ser expressa através do RNA mensageiro, formando uma seqüência de aminoácidos característica de um peptídeo específico. O termo "gene" inclui regiões não codificantes (regulatórias), as quais não são transcritas, bem como regiões codificantes para RNAs que não serão posteriormente traduzidas na forma de proteínas.

25 A seqüência da presente invenção foi obtida a partir do DNA extraído de células humanas.

**"Oligonucleotídeos"** refere-se a um ácido nucléico de fita simples com tamanho variando entre 2 e 60 bases. Os oligonucleotídeos freqüentemente são sintéticos, mas podem

também ser produzidos a partir de polinucleotídeos que ocorrem naturalmente. Uma sonda é um oligonucleotídeo capaz de se ligar a um ácido nucléico alvo com seqüência complementar através de um ou mais tipos de ligações químicas, usualmente  
5 através do pareamento complementar de bases via pontes de hidrogênio. As sondas oligonucleotídicas possuem freqüentemente de 5 a 60 nucleotídeos, variando mais freqüentemente entre 15 e 30 bases de tamanho. Um oligonucleotídeo pode incluir bases naturais (A, T, G, C) ou  
10 modificadas (7-deazaguanosina, iosina, etc). Além disto, as bases podem ser unidas por ligações outras que não a ligação fosfodiéster, tais como uma ligação fosfodiéster ou uma ligação fosforotioato.

**"Farmacogenômica e Farmacogenética"**, Farmacogenômica é a  
15 aplicação de abordagens e tecnologias empregadas na Genômica e identificação de alvos para drogas. Em outras palavras, a farmacogenômica utiliza informação genética para predizer se uma droga irá contribuir positivamente ou negativamente no tratamento de um paciente. Farmacogenética é o estudo da  
20 variação da ação e reação à drogas e sua relação com o conteúdo genético do paciente. Os indivíduos respondem diferentemente aos tratamentos com drogas; alguns positivamente, outros sem modificações óbvias em sua condição e ainda outros com efeitos colaterais e/ou reações alérgicas.  
25 A Farmacogenética é um subconjunto da farmacogenômica, empregando métodos e técnicas da Genômica e da Bioinformática para a identificação de correlações entre características genômicas como, por exemplo, SNPs e perfis de resposta característicos.

**"Polimorfismos"**, em termos gerais, refere-se à habilidade de algo aparecer em várias formas. Especificamente, a existência de duas ou mais formas alternativas (alelos) de um loco cromossomal diferindo na seqüência nucleotídica.

5 **"Seqüência referência"** é a seqüência nucleotídica do gene que codifica para a enzima NAT2 (SEQ ID NO:1) e a seqüência do aminoácido correspondente da proteína NAT2 (SEQ ID NO:2), como descrito por Rieder et al 2003 (submissão direta) número de acesso no geneBank (AY331807).

10 **"Polimorfismo de base única ou SNP"** refere-se à substituição de um único nucleotídeo por outro com os dois ou mais estados sendo encontrados na população.

**"Hospedeiro"** inclui procariotos e eucariotos, tais como, bactérias, leveduras e fungos filamentosos, bem como células vegetais e animais. O termo inclui um organismo ou célula que  
15 recebe um veículo de expressão ou de clonagem replicável.

**"Promotor"** refere-se à seqüência de DNA reconhecida por uma RNA polimerase. A presença de tal seqüência permite que a RNA polimerase se ligue ao DNA e inicie a transcrição de  
20 seqüências gênicas operacionalmente ligadas.

**"Região promotora"** é a região que inclui a seqüência promotora ou regulatória, bem como outras seqüências, as quais, podem ser necessárias para a iniciação da transcrição. A presença da região promotora é necessária para levar a  
25 expressão de outras seqüências gênicas operacionalmente ligadas.

**"Fenótipo de acetilação rápida"** refere-se as características de metabolização de compostos aminados, que em

alguns indivíduos é rápida em comparação com outros indivíduos.

**"Fenótipo de acetilação lenta"** é a característica de um indivíduo, no qual, a acetilação de compostos aminados é lenta  
5 em comparação com outros indivíduos.

**"Reações adversas"**, referem-se a efeitos não esperados ou negativos induzidos por uma droga durante um tratamento quimioterápico. As reações adversas podem ser locais, sistêmicas ou alergênicas.

## 10 **Descrição Detalhada da Invenção**

Esta invenção relata a identificação de seis novos SNPs na região codificante do gene que codifica para a proteína N-acetiltransferase 2 (NAT2) humana em indivíduos residentes em duas diferentes regiões geográficas do Brasil. Os SNPs aqui  
15 descritos foram identificados através das técnicas de Reação em Cadeia da Polimerase (PCR) e seqüenciamento dos produtos amplificados. A detecção de polimorfismos em seqüências específicas de DNA pode ser obtida por uma variedade de métodos incluindo, mas não limitada a i) polimorfismo de  
20 tamanho de fragmentos de restrição (PCR-RFLP) baseado na clivagem específica por uma enzima de restrição de um sitio alelo específico, (kan e Dozy, Lancet ii:910-912, 1978); ii) hibridação com sondas oligonucleotídicas alelo específicas (Wallace e cols. Nucl Acids Res. 6:3543-3557 1978), incluindo  
25 oligonucleotídeos imobilizados (Saiki e cols, Proc. Natl. Acad. Sd. USA 86:6230-6234 1969) ou arranjos de oligonucleotídeos (Maskos na Southern Nucl. Acids Res. 17:2503-2516 1993); iii) PCR alelo específico (Newton e cols., Nucl Acids Res 17:2503-2516 1989); iv) detecção e reparo de



pareamento errôneo "mismatch-repair detection" (MRD) (Faham & Cox, *Genome Res.* 5:474-482, 1995); v) gel de eletroforese desnaturante de gradiente "Denaturing-gradient gel electrophoresis" (DGGE) (Fisher e cols, *Proc. Natl. Acad. Sci. USA* 80:1579-1583 1983); vi) detecção de polimorfismo por conformação de fita simples "Single strand conformation polymorphism" (SSCP) (Orita e cols, *Genomics* 5:874-879 1983); vii) clivagem de pareamento errôneo por RNase (Meyers e cols, *Science* 230:1242 1985); viii) método químico (Cotton e cols, *Proc. Natl. Acad. Sci. USA* 85:4397-440 1988) ou enzimático (Youil e cols, *Proc. Natl. Acad. Sci. USA* 92:87-91 1995); ix) clivagem de DNA heteroduplex; x) métodos baseados na extensão de primers alelo específicos (Syvanen e cols, *Genomics* 8:684-692 1990); xi) ensaio de ligação de oligonucleotídeo "Oligonucleotide ligation Assay" (OLA) (Landegren e cols, *Science* 241:1077 1988); xii) reação em cadeia de ligação alelo específica "allele-specific ligation chain reaction (LCR) (Barrany, *Proc. Natl. Acad. Sci. USA* 88:189-193 1991); xiii) sequenciamento radioativo ou fluorescente (Sambrook, J. & Russell, D. W. 2001; *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Os 6 polimorfismos descritos nas FIGs. 1A-1B e Tabela 2 estão localizados na região codificante do gene *NAT2* e incluem as seguintes substituições (T → C na base 29, G → T na base 152, G → A na base 203, C → T na base 228, C → T na base 458, e A → G na base 600). Com exceção dos SNPs C228T e A600G que são silenciosos, todos os outros 4 SNPs modificam o aminoácido transcrito nas seguintes posições (SNP T 29 C troca de I → T

no aminoácido 10, SNP G 152 T, troca de G → V no aminoácido 51, SNP G 203 A, troca de C → Y no aminoácido 68, SNP C 458 T, troca de T → I no aminoácido 153).

Como descrito acima, a presente invenção está relacionada a uma seqüência de DNA compreendida na região codificante do gene *NAT2*, na qual foram identificados 6 novos SNPs. Variantes muito próximas também estão incluídas como parte desta invenção, bem como ácidos nucléicos recombinantes que contenham qualquer dos SNPs aqui descritos ou as regiões acima ou abaixo dos mesmos na seqüência descrita nas FIGs. 1A-1B e reivindicações.

Esta invenção também está relacionada à utilização de métodos que utilizem ácidos nucléicos (DNA ou RNA) isolados e/ou recombinantes que sejam caracterizados pela sua capacidade de hibridarem a um ácido nucléico que codifique uma proteína ou polipeptídeo, tal como um ácido nucléico que contenha a seqüência descrita nas FIGs. 1A-1B ou parte de um DNA estranho que contenha o mínimo de nucleotídeos suficientes para codificar um polipeptídeo que contenha parte da seqüência de aminoácidos descrita nas FIGs. 1A-1B e Tabela 2 ou que codifiquem para equivalentes funcionais, ou seja, um polipeptídeo que quando incorporado a uma célula hospedeira tenha toda ou parte da atividade da proteína *NAT2*. Um equivalente funcional da proteína *NAT2* deve ter uma seqüência de aminoácido similar (no mínimo 65% de identidade) e características similares a proteína *NAT2*. Um ácido nucléico que hibrida com outro ácido nucléico que codifica para a proteína *NAT2* ou polipeptídeo, tais como descrito nas FIGs. 1A-1B e Tabela 2 podem ser de fita dupla ou de fita simples.

Hibridação com um DNA contendo toda ou parte da seqüência descrita nas FIGs. 1A-1B inclui hibridação tanto com a fita mostrada como com a sua fita complementar.

Ácidos nucleicos incluindo DNA ou RNA podem ser detectados e isolados por hibridação sob condições de estringência alta ou moderada, as quais são escolhidas de forma a não permitir a hibridação de seqüências nucleotídicas não complementares. "Condições de alta estringência" e "condições de estringência moderada" para a hibridação de ácidos nucleicos são exemplificadas em Sambrook, J. & Russell, D. W. 2001; *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. As condições exatas que determinam a estringência da hibridação dependem não somente da força iônica, temperatura e concentração de agentes desestabilizadores, tais como formamida, mais também de fatores tais como o tamanho da seqüência de ácido nucleico, a composição das bases, percentual de pareamento errôneo entre as duas seqüências e a freqüência de ocorrência. Assim, as condições de alta ou moderada estringência podem ser calculadas empiricamente.

A presente invenção também diz respeito ao uso da seqüência nucleotídica aqui descrita, ou parte da mesma, para a identificação de sondas de DNA específicas para o gene *NAT2* ou oligonucleotídeos para a amplificação (através de PCR) do gene *NAT2*.

A presente invenção também descreve, na Tabela 1, oligonucleotídeos para a amplificação de dois fragmentos, do gene de *NAT2*, que, juntos, compreendem toda a região codificante do respectivo gene contendo os polimorfismos

descritos. Os seis novos SNPs presentes na região codificante do gene *NAT2* foram identificados a partir do seqüenciamento automático dos fragmentos gerados após a amplificação do DNA por PCR com os dois pares de oligonucleotídeos descritos na Tabela 1. Assim, estes oligonucleotídeos podem ser utilizados para detecção dos mesmos SNPs em quaisquer amostras de DNA provenientes de material biológico humano através do PCR e seqüenciamento.

Os seis novos SNPs descritos nas FIGs. 1A-1B e tabela 2 da presente invenção podem ser utilizados para a determinação de haplótipos, os quais, quando estendidos e associados aos SNPs já descritos, podem ser utilizados para a predição do fenótipo de acetilação (acetilador lento ou rápido) de um determinado indivíduo. Como descrito no exemplo 1, o DNA pode ser isolado de diferentes indivíduos e, através de técnicas de seqüenciamento de DNA, a sequência do gene de *NAT2* de um indivíduo pode ser obtida. Após a identificação dos polimorfismos, o genótipo é determinado permitindo assim a predição do fenótipo.

A invenção também está relacionada a um método de criação de um protocolo prognóstico para um paciente recebendo algum agente terapêutico metabolizado por *NAT2* tais como, amonafida, isoniazida, fenilzina, hidrazina, dapsona, procainamida, sulfametazina e outras sulfonamidas. O método inclui: a) identificação de pacientes recebendo uma destas drogas; b) determinação do fenótipo de acetilação (rápidos ou lentos) com base no genótipo; e c) conversão dos dados obtidos no passo (b) em um protocolo prognóstico. O protocolo prognóstico pode incluir a prevenção da eficácia da droga, predição do

prognóstico do paciente, predição de interações entre drogas, e predição de efeitos adversos. A combinação das seqüências de DNA geradas com o fenótipo de acetilação fornece um protocolo prognóstico específico para àquele indivíduo. Por exemplo, 5 estudos têm mostrado que a isoniazida é mais tóxica em pacientes com fenótipo de acetiladores lentos que em pacientes com fenótipo de acetiladores rápidos. Portanto, a identificação destes pacientes usando as seqüências nucleotídicas fornecidas aqui pode auxiliar no desenho de 10 esquemas terapêuticos balanceando eficácia e toxicidade.

Em outro exemplo, relacionado a um protocolo prognóstico, pacientes identificados como acetiladores lentos estão sob risco de hipersensibilidade cutânea quando tratados com trimetoprin-sulfametoxazole (TMP-SMZ). Desta forma, antes de 15 prescrever uma droga particular, tal como trimetoprin-sulfametoxazole (TMP-SMZ), o médico pode solicitar um teste para a identificação do fenótipo de acetilação do paciente baseado no perfil genotípico e, uma vez o paciente sendo um acetilador lento o médico então pode prescrever uma 20 terapêutica alternativa.

### **Ensaaios Clínicos**

Esta invenção também relaciona-se a um método para auxiliar na avaliação de novos fármacos através de ensaios clínicos (*Clinical Trials*). O método inclui: a) administração 25 de um agente terapêutico (metabolizado por NAT2) a um indivíduo e mensuração de sua eficácia; b) determinação do status de acetilação do indivíduo (rápido ou lento) baseado no genótipo ou haplótipo do gene de NAT2 incluindo os SNPs descritos na presente invenção e c) determinação a partir dos

itens (a) e (b) do agente terapêutico seria mais eficaz para aquele genótipo em particular. Os ensaios clínicos geralmente baseiam-se em informações fornecidas pelo paciente, incluindo idade, sexo e histórico familiar. A invenção fornece uma  
5 seqüência de DNA do gene de NAT2 contendo novos SNPs os quais podem ser adicionados ao banco de SNPs e utilizados como fatores genéticos do paciente em um ensaio clínico. Como descrito aqui, um genótipo individual pode ser determinado por métodos de seqüenciamento de DNA descritos no exemplo 1 desde  
10 que seus haplótipos sejam de fase conhecida.

Após a administração da droga o genótipo do paciente pode então ser comparado com a eficácia da mesma e ocorrência de quaisquer efeitos adversos. Baseado nesta informação, fármacos podem ser desenvolvidos especificamente para genótipos  
15 individuais os quais estejam relacionados com uma maior eficácia e/ou menor toxicidade. Genótipos de pacientes que não respondem ao fármaco podem ser agrupados e fármacos que utilizem vias alternativas de acetilação podem ser desenvolvidos.

## 20 **Frequência dos dados**

A presente invenção também relata a frequência dos novos SNPs em duas diferentes populações do Brasil, como descrito na Tabela 3. Os dados mostram que os SNPs identificados aparecem em diferentes frequências e são específicos nas duas  
25 populações estudadas, o que corrobora a hipótese de se tratarem de duas populações geneticamente distintas e conseqüentemente com fenótipos de acetilação também distintos. Os SNPs T29C, G152T, G203A e C458T ocorrem exclusivamente na

amostra populacional do Rio de Janeiro e os SNPs C228T e A600G ocorrem exclusivamente na amostra populacional de Goiás.

Como descrito anteriormente, a enzima NAT2 tem sido associada a uma variedade de doenças e desordens incluindo 5 câncer de bexiga, câncer de colo, câncer de próstata, carcinoma celular urotelial e doença de Gilbert's. Mais particularmente, a identificação de indivíduos mais suscetíveis a metabolização de compostos com potencial mutagênico e/ou carcinogênico incluindo 2-aminofluoreno, 4- 10 aminobifenil, benzidina, beta-naftilamina e certas arilaminas heterocíclicas presentes em pirolisados de proteínas, pode ser muito benéfica na prevenção, por parte do indivíduo, da utilização de tais compostos. Os inventores disponibilizam uma seqüência de DNA contendo novos SNPs, os quais podem ser 15 úteis para a identificação de indivíduos, com polimorfismos no gene de NAT2 que estão associados com estas doenças e afetam o metabolismo dos compostos descritos acima.

Ácidos nucleicos associados a doenças e úteis para métodos diagnósticos incluem oligonucleotídeos com um tamanho mínimo 20 de 15 nucleotídeos, preferencialmente um tamanho mínimo de 20 nucleotídeos e mais preferencialmente um tamanho mínimo de 25 nucleotídeos que hibride especificamente com um ácido nucleico associado a doenças.

A amostra a ser analisada deve ser submetida a um processo 25 de extração de DNA. Entende-se que o método de extração de DNA dependerá da natureza da amostra a ser analisada. O DNA resultante deverá ser submetido a PCR para amplificação da região contendo os SNPs e em seguida seqüenciados.

A seqüência polinucleotídica descrita na presente invenção (ou parte da mesma), principalmente aquela relacionada a identificação dos SNPs aqui descritos, representa uma valiosa fonte de informação para a caracterização de indivíduos em termos de haplótipos e outros subgrupos, como por exemplo na 5 investigação de suscetibilidade ao tratamento com alguns fármacos particulares. Esta abordagem é ainda facilitada pelo armazenamento destas informações em uma base de dados informatizada e sua posterior análise por programas de 10 bioinformática. Assim, a seqüência polinucleotídica da presente invenção é particularmente útil como componente de banco de dados utilizado para a identificação de seqüências e outras buscas. O estoque das informações da seqüência em qualquer mídia e sua utilização em bancos de dados de 15 seqüências com relação a "polinucleotídeo ou seqüência de polinucleotídeos da invenção" cobre qualquer característica química ou física do polinucleotídeo da presente invenção que pode ser reduzido a, convertido em, ou estocado em alguma mídia, tais como, um disquete ou um CD, preferencialmente numa 20 forma legível em um computador.

### **Exemplo 1**

Amostras de sangue foram coletadas de 404 indivíduos brasileiros consanguineamente não relacionados para a genotipagem de NAT2. As amostras utilizadas compreenderam 25 indivíduos de duas regiões geograficamente distintas do Brasil (Rio de Janeiro e Goiás) e foram coletadas no Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (UFRJ) e no Hospital de Referência Anuar Auad do Estado de Goiás. O DNA genômico foi isolado deste



material utilizando o kit comercial FlexiGene DNA Kit (Qiagen Inc., USA)

Os fragmentos de DNA cobrindo toda a região codificante do gene *NAT2* para posterior seqüenciamento foram gerados a partir da amplificação pela técnica de "reação de polimerase em cadeia" (PCR), através da utilização de um par de oligonuclçotídeos (NAT2ED e NAT2ER, Tabela 1). As condições para a reação de PCR foram as seguintes: 100 ng de DNA genômico, 200 ng de cada oligonucleotídeo, 0,2 mM de dNTPs, 4.0 mM MgCl<sub>2</sub>, 10% glicerol (Invitrogen Life Technologies, Gaithersburg, MD, USA), 10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% Triton X-100, and 1.1 U da mistura de duas DNA polymerases (1 U of *Taq* DNA polymerase and 0.1 U de *Tli* DNA polymerase - Promega Corporation, USA). Foram usados em um volume final de 50 µl para cada amostra. A amplificação foi realizada nas seguintes condições de ciclagem: uma desnaturação inicial a 94°C for 5 min, seguida de 35 ciclos de 94°C por 1 min. 57.5°C for 1 min e 72°C por 1 minuto. O passo de extensão final foi realizado a 72°C por 5 min. Após a amplificação, aproximadamente 20 % do produto amplificado foi examinado em gel de agarose. Antes da etapa de sequenciamento do DNA, os produtos de PCR foram purificados com acetato de amônio 2.5 M.

**Tabela 1**

Seqüência dos oligonucleotídeos utilizados para a amplificação e/ou sequenciamento da região codificante do gene *NAT2*

NAT2 ED	5` - TTA GTC ACA CGA GGA AAT CAA A - 3`
NAT2 ER	5` - AAA TGC TGA CAT TTT TAT GGA TGA - 3`
NAT2 ID	5` - ACC ATT GAC GGC AGG AAT TA - 3`
NAT2 IR	5` - TGG TCC AGG TAC CAG ATT CC - 3`

<sup>ED</sup>- Externo direto; <sup>ER</sup>- Externo Reverso; <sup>ID</sup>- Interno direto; <sup>IR</sup>- Interno Reverso

Quatro reações de seqüenciamento foram realizadas para cada amostra utilizando o Kit ABI PRISM Big Dye Terminator v. 3.1 (PE Applied BioSystems), de acordo com as recomendações do fabricante em um seqüenciador de capilar "ABI PRISM 3730 DNA  
5 Analyzer" (PE Applied BioSystems). Em cada reação (10 µl de volume final), foram utilizados aproximadamente 20 ng de produto amplificado, 3,2 pmol do iniciador, 0,5 µl de Big Dye e 1,5 µl de tampão 5X (fornecido no kit).

A análise das seqüências e identificação das SNPs  
10 presentes na região codificante de *NAT2* foi feita através do software SeqScape v. 2.5 (Applied Biosystems). (<http://www.appliedbiosystems.com/catalog/myab/StoreCatalog/products/CategoryDetails.jsp?hierarchyID=101&category3rd=111917&trail=no>). A seqüência AY331807  
15 (<http://www.ncbi.nlm.nih.gov/GenBank>) foi usada como referência. As seqüências obtidas, referentes a cada amostra clínica individual, foram agrupadas sob forma de projeto de acordo com a utilização do SeqScape. As seqüências foram comparadas por alinhamento com a seqüência referência e os  
20 sítios polimórficos (SNPs), foram identificados.

As figuras 1A-1B mostram as seqüências nucleotídica e aminoacídica do tipo selvagem do gene *NAT2* descrita por Blum et al. (*DNA and Cell Bio.*, 9:192-203, 1990), incluindo o sítio de iniciação da tradução "ATG". A posição dos seis novos SNPs  
25 descobertos estão sublinhadas e em negrito. Adicionalmente, as trocas aminoacídicas também estão sublinhadas e em negrito.

A Tabela 2 contém a lista dos polimorfismos de base única (SNPs) descobertos no gene *NAT2*. A posição do nucleotídeo de acordo com as FIGs. 1A-1B é listada na primeira coluna. A

segunda coluna descreve as trocas de bases a partir do genótipo selvagem. A posição do aminoácido afetado pela troca da base nucleotídica é listada na terceira coluna. As trocas de aminoácidos decorrentes de 4 dos 6 SNPs adicionais (base 29, 152, 203, 228, 458 e 600) são listadas na quarta coluna.

**Tabela 2**

Posição dos SNPs	Troca de nucleotídeo	posição do aminoácido	Troca de aminoácido
29	T para C	10	I para T
152	G para T	51	G para V
203	G para A	68	C para Y
228	C para T	76	Não ocorre
458	C para T	153	T para I
600	A para G	158	Não ocorre

### **Significância dos novos SNPs**

Nenhuma troca de aminoácido ocorre quando da presença do SNP na base 228 e 600 da sequência descrita nas FIGs. 1A-1B. Trata-se de SNPs localizados no terceiro nucleotídeo dos códons 76 e 158 respectivamente.

Esta invenção fornece ainda quatro SNPs adicionais, não sinônimos, presentes na região codificante do gene de NAT2, os quais alteram a sequência de aminoácidos da proteína.

A substituição no nucleotídeo 29 (T→C) na região codificante do gene de NAT2, resulta numa troca de aminoácido No códon 10 (I→T: isoleucina para treonina). Neste caso temos a troca de um aminoácido cadeia lateral apolar, hidrofóbico por um polar. Tendo em vista a diferença entre as cadeias laterais destes aminoácidos, muito provavelmente esta substituição pode afetar a estrutura, flexibilidade, estabilidade e enovelamento da proteína NAT2.

A substituição no nucleotídeo 152 (G→T) na região codificante do gene de NAT2, resulta numa troca de aminoácido no códon 51 (G→V: glicina para valina). Aparentemente uma troca conservativa (aminoácidos com cadeias laterais similares). Contudo, se localizada em alguma região crítica (sítios de ligação do substrato ou co-fatores), esta troca pode afetar a estrutura ou atividade da proteína.

A substituição no nucleotídeo 203 (G→A) na região codificante do gene de NAT2 resulta numa troca de aminoácido na posição 68 (C→Y: cisteína para tirosina). Neste caso temos a troca de um aminoácido apolar por um polar não carregado. Tendo em vista a diferença entre as cadeias laterais destes aminoácidos, é provável que esta substituição afete a estrutura, flexibilidade, estabilidade e enovelamento da proteína NAT2.

A substituição no nucleotídeo 458 (C→T) na região codificante do gene de NAT2, resulta numa troca de aminoácido na posição 153 (T→I: treonina para isoleucina). Neste caso temos a troca de um aminoácido polar não carregado por um apolar. Tendo em vista a diferença entre as cadeias laterais destes aminoácidos, muito provavelmente esta substituição pode afetar a estrutura, flexibilidade, estabilidade e enovelamento da proteína NAT2.

A combinação de um ou todos os SNPs aqui descobertos com aqueles já descritos na literatura pode ter efeitos adicionais na estrutura, atividade, dobramento e estabilidade da proteína. Evidências sugerem que o fenótipo metabólico seja gerado a partir da combinação de todos os SNPs presentes no indivíduo. Utilizando análises de regressão linear múltipla,

pesquisadores foram capazes de formular expressões matemáticas capazes de prever a capacidade metabólica da enzima NAT2 baseados no genótipo (Maisel et al., Pharmacogenomics, 1997). De acordo com estas análises, todas as substituições de nucleotídeos, mesmo aquelas que não causam a troca de aminoácidos, afetam o fenótipo de alguma forma. Embora eles possam prever com razoável segurança o fenótipo da maioria dos indivíduos, se todos os SNPs conhecidos forem levados em consideração a precisão do modelo é consideravelmente maior. Uma vez que neste modelo os autores tiveram um indivíduo para o qual o fenótipo não pode ser previsto, existe uma alta probabilidade da existência de SNPs adicionais no gene cujas análises realizadas não foram capazes de detectar.

#### **Dados de Frequência**

A Tabela 3 mostra a frequência alélica dos seis novos SNPs descritos na presente invenção, em duas regiões distintas do Brasil. A primeira coluna lista os diferentes SNPs novos identificados nessa amostragem. A segunda, quinta e oitava colunas indicam as frequências dos SNPs na amostra total, no Rio de Janeiro e em Goiás, respectivamente. As colunas 3, 4, 6, 7, 9 e 10 detalham se e quantos indivíduos foram homocigotos ou heterocigotos para o polimorfismo listado. A frequência na segunda coluna refere-se ao número de alelos com polimorfismo dividido pelo número total de alelos na amostragem. A troca de bases e posição referem-se as coordenadas das FIGS. 1A-1B.

A Tabela 4 lista a frequência dos SNPs descritos na invenção com os sítios polimórficos adicionais, suas respectivas frequências nas diferentes regiões estudadas e,

para cada SNP, se a variante mutante foi encontrada em homozigose ou heterozigose. A frequência é o número de alelos contendo o polimorfismo dividido pelo número total de alelos na amostragem. A troca de bases e a posição das bases, 5 referem-se as coordenadas das FIGS. 1A-1B.

**Tabela 3**

**Frequência alélica dos SNPs da presente invenção no gene *NAT2* em duas diferentes regiões do Brasil**

SNPs	TOTAL			RIO DE JANEIRO			GOIÁS		
	%	Heterozigotos	Homozigotos	%	Heterozigotos	Homozigotos	%	Heterozigotos	Homozigotos
<b>29C</b>	0,12	01/404	0	0,17	01/298	0	0	0	0
<b>152T</b>	0,12	01/404	0	0,17	01/298	0	0	0	0
<b>203 A</b>	0,12	01/404	0	0,17	01/298	0	0	0	0
<b>228 T</b>	0,12	01/404	0	0	0	0	0,47	01/106	0
<b>458 T</b>	0,12	01/404	0	0,17	01/298	0	0	0	0
<b>600G</b>	0,12	01/404	0	0	0	0	0,47	01/106	0

**Tabela 4**

**Frequência dos SNPs identificados no gene *NAT2* em duas diferentes regiões do Brasil**

SNPs	TOTAL			RIO DE JANEIRO			GOIÁS		
	%	Heterozigotos	Homozigotos	%	Heterozigotos	Homozigotos	%	Heterozigotos	Homozigotos
<b>29C</b>	0.1	01/404	0	0.1	01/298	0	0	0	0
<b>70A</b>	0.1	01/404	0	0.2	01/298	0	0	0	0
<b>152T</b>	0.1	01/404	0	0.2	01/298	0	0	0	0
<b>191A*</b>	4.1	31/404	01/404	5	26/298	01/298	2	05/106	0
<b>203A</b>	0.1	01/404	0	0.2	01/298	0	0	0	0
<b>228T</b>	0.1	01/404	0	0	0	0	0.5	01/106	0

<b>282T*</b>	37	184/404	57/404	38	144/298	41/298	34	40/106	16/106
<b>341C*</b>	38	176/404	65/404	40	136/298	50/298	33	40/106	15/106
<b>345T</b>	0.3	02/404	0	0	0	0	1	02/106	0
<b>403G</b>	0.6	04/404	0	1	04/298	0	0.5	0	0
<b>458T</b>	0.1	01/404	0	0.2	01/298	0	0	0	0
<b>472C</b>	0.6	05/404	0	0.5	03/298	0	1	02/106	0
<b>481T*</b>	35	176/404	54/404	36	137/298	40/298	32	39/106	14/106
<b>590A*</b>	27	157/404	29/404	26	119/298	18/298	28	38/106	11/106
<b>600G</b>	0.1	01/404	0	0	0	0	0.5	01/106	0
<b>609T</b>	0.4	03/404	0	0.3	02/298	0	0.5	01/106	0
<b>683T</b>	0.2	02/404	0	0.3	02/298	0	0	0	0
<b>766G</b>	0.7	06/404	0	0.7	04/298	0	1	02/106	0
<b>803G*</b>	40	184/404	69/404	41	143/298	51/298	36	41/106	18/106
<b>838A</b>	0.5	04/404	0	0.5	03/298	0	0.5	01/106	0
<b>G857A*</b>	4.4	36/404	0	4.5	27/298	0	4.2	09/106	0

Apesar de ilustrada e descrita aqui com referência a determinadas representações específicas, a presente invenção não tem, no entanto, a intenção de estar limitada aos detalhes mostrados. Várias modificações podem ser feitas nos detalhes dentro do âmbito e alcance de equivalentes sem se afastar do espírito da invenção.

-50

TGCTAAAGTATGATATGTTTTATGTTTTGTTTTCTTGCTTAGGGGATCATGGACATTG  
M D I

11

**T29C**

AAGCATATTTTGAAGAATTTGGCTATAAGAACTCTAGGAACAAATTGGACTTGAAACAT  
E A Y F E R I G Y K N S R N K L D L E T

71

TAACTGACATTCTTGAGCACCAGATCCGGGCTGTTCCCTTTGAGAACCTTAACATGCATT  
L T D I L E H Q I R A V P F E N L N M H

131

**G152T**

GTGGGCAAGCCATGGAGTTGGCTTAGAGGCTATTTTTGATCACATTGTAAGAAGAAACC  
C G Q A M E L G L E A I F D H I V R R N

191

**G203A**

**C228T**

GGGGTGGGTGGTGTCTCCAGGTCAATCAACTTCTGTACTGGGCTCTGACCACAATCGGTT  
R G G W C L Q V N Q L L Y W A L T T I G

251

TTCAGACCACAATGTTAGGAGGGTATTTTTACATCCCTCCAGTTAACAAATACAGCACTG  
F Q T T M L G G Y F Y I P P V N K Y S T

311

GCATGGTTCACCTTCTCCTGCAGGTGACCATTGACGGCAGGAATTACATTGTTCGATGCTG  
G M V H L L L Q V T I D G R N Y I V D A

371

GGTCTGGAAGCTCCTCCCAGATGTGGCAGCCTCTAGAATTAATTTCTGGGAAGGATCAGC  
G S G S S S Q M W Q P L E L I S G K D Q

431

**C458T**

CTCAGGTGCCTTGCATTTTCTGCTTGAAGAAGAGAGAGGAATCTGGTACCTGGACCAA  
P Q V P C I F C L T E E R G I W Y L D Q

491

TCAGGAGAGAGCAGTATATTACAAACAAAGAATTTCTTAATTCTCATCTCCTGCCAAAGA  
I R R E Q Y I T N K E F L N S H L L P K

551

**A600G**

AGAAACACCAAAAAATATACTTATTTACGCTTGAACCTCGAACAATTGAAGATTTTGAGT  
K K H Q K I Y L F T L E P R T I E D F E

**FIGURA 1A**



611  
CTATGAATACATACCTGCAGACGTCTCCAACATCTTCATTTATAACCACATCATTTTGT  
S M N T Y L Q T S P T S S F I T T S F C

671  
CCTTGCAGACCCCAGAAGGGGTTTACTGTTTGGTGGGCTTCATCCTCACCTATAGAAAAT  
S L Q T P E G V Y C L V G F I L T Y R K

731  
TCAATTATAAAGACAATACAGATCTGGTCGAGTTTAAACTCTCACTGAGGAAGAGGTTG  
F N Y K D N T D L V E F K T L T E E E V

791  
AAGAAGTGCTGAAAAATATATTTAAGATTTTCCTTGGGGAGAAATCTCGTGCCCAAACCTG  
E E V L K N I F K I S L G R N L V P K P

851  
GTGATGGATCCCTTACTATTTAG  
G D G S L T I \*

**FIGURA 1B**

## REIVINDICAÇÕES

1. Ácido nucléico isolado compreendendo pelo menos 15 bases nucleotídicas incluindo o sítio polimórfico caracterizado por ser selecionado de um grupo consistindo de:

- 5 (a) uma substituição T → C no nucleotídeo 29 da SEQ ID NO:1
- (b) uma substituição G → T no nucleotídeo 152 da SEQ ID NO:1
- (c) uma substituição G → A no nucleotídeo 203 da SEQ ID NO:1
- (d) uma substituição C → T no nucleotídeo 228 da SEQ ID NO:1
- (e) uma substituição C → T no nucleotídeo 458 da SEQ ID NO:1
- 10 (f) uma substituição A → G no nucleotídeo 600 da SEQ ID NO:1.

2. Ácido nucleico isolado de acordo com a reivindicação 1 caracterizado por compreender um DNA.

3. Sonda específica para um alelo isolado caracterizado por ser capaz de detectar um sítio polimórfico da SEQ ID NO:1 da reivindicação 1.

4. Oligonucleotídeo específico para um alelo isolado caracterizado por ser capaz de detectar um sítio polimórfico da SEQ ID NO:1 da reivindicação 1.

5. Kit diagnóstico caracterizado por utilizar a sonda alelo específica conforme reivindicação 3 ou um oligonucleotídeo alelo específico conforme a reivindicação 4.

6. Método de identificação de qualquer polimorfismo da SEQ ID NO:1 em mamíferos caracterizado por compreender as seguintes etapas:

25 (a) preparo de amostras de células ou tecidos de mamíferos;

(b) identificação de tecidos ou células contendo todos ou parte dos polimorfismos da SEQ ID NO:1 da reivindicação 1, em condições onde o DNA hibridado possa ser produzido;

(c) identificação do DNA hibridado;

(d) clonagem e seqüenciamento do DNA hibridado para a identificação do gene *NAT2* em mamíferos.

7. Método de identificação individual em um ensaio clínico  
5 caracterizado por compreender:

(a) produção de uma biblioteca de SNPs incluindo os sítios polimórficos da SEQ ID NO:1 conforme a reivindicação 1 e seu respectivo fenótipo;

(b) seqüenciamento individual do gene *NAT2*;

10 (c) conjugação do genótipo obtido no Ítem (b) com o fenótipo obtido em (a).

8. Método de identificação de composições terapêuticas as quais são eficazes em indivíduos caracterizado por compreender:

15 (a) administração de composição terapêutica para um indivíduo e mensuração de sua eficácia;

(b) determinação de genótipos individuais e de sítios polimórficos da SEQ ID NO:1 da reivindicação 1 que indiquem se o indivíduo é um acetilador rápido ou lento;

20 (c) determinação a partir das etapas (a) e (b) de quais composições terapêuticas serão as mais efetivas para um genótipo particular e quais terão efeitos adversos.



## CAPÍTULO III

Essa etapa incluiu descrição de novos polimorfismos no gene *NAT2* bem como a caracterização de novos alelos por clonagem e re-sequenciamento e experimentos de modelagem molecular onde verificamos a influência de diferentes SNPs na estrutura protéica de *NAT2*. Esses resultados encontram-se reunidos no manuscrito “Sequence analysis of *NAT2* gene of Brazilians: identification of undescribed Single Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure” **submetido** para a revista internacional *Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis* em abril de 2009.

Elsevier Editorial System(tm) for Mutation Research/Fundamental and Molecular  
Mechanisms of Mutagenesis

Manuscript Draft

Manuscript Number: MUT-D-09-00087

Title: Sequence analysis of NAT2 gene of Brazilians: identification of undescribed Single  
Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure

Article Type: Research Paper

Keywords: human NAT2, polymorphism, molecular modeling, N-acetylation, structure/function,  
Brazilian population

Corresponding Author: Doctor Adalberto Rezende Santos, Ph.D

Corresponding Author's Institution: Oswaldo Cruz Foundation

First Author: Raquel F Teixeira, MSc

Order of Authors: Raquel F Teixeira, MSc; Floriano P Silva Junior, Ph.D; Aline R Silveira, MSc;  
Pedro H Cabello, Ph.D; Leila Mendonça-Lima, Ph.D; Marcelo F Rabahi, Ph.D; Afrânio L Kritski,  
Ph.D; Fernanda Q Mello, Ph.D; Philip N Suffys, Ph.D; Antonio B Miranda, Ph.D; Adalberto  
Rezende Santos, Ph.D

Abstract: N-acetyltransferase 2 (NAT2) metabolizes a variety of xenobiotics that include many  
therapeutic drugs, chemicals and carcinogens. This enzyme is genetically variable in human  
populations and polymorphisms in the NAT2 gene have been associated with drug toxicity and  
efficacy as well as cancer susceptibility. Here we have focused on the identification of NAT2  
variants in Brazilian individuals resident in two different regions by direct sequencing and in the  
characterization of new haplotypes after cloning and re-sequencing. Upon analysis of DNA  
samples from 404 individuals, six new mutations (c.29T>C, c.152G>T, c.203G>A, c.228C>T,  
c.458C>T and c.600A>G) and seven new NAT2 alleles were identified with different frequencies in

Rio de Janeiro and Goiás. Molecular modeling and structural analysis suggested that polymorphism p.Gly51Val may have an important effect on substrate recognition by NAT2. We also observed that mutation p.Cys68Tyr would impair acetylating activity due to the resulting geometric restrictions and incompatibility of the functional group in the Tyr side chain with the admitted chemical mechanism for catalysis by NATs. Moreover, other variants may lead to the presence of hydrophobic residues involved in protein aggregation and/or targeted degradation. From the new alleles, NAT2\*5N, \*5O, \*6H, \*6M, and \*12E may code for a slow activity and NAT2\*20 and \*12I, are functional alleles. Finally, it is possible that NAT2\*12F and NAT2\*12J do not interfere with the acetylation activity because they contain other mutations (c.622T>C and c.29T>C, in addition to c.803A>G) which do not affect the enzyme structure. Functional studies are needed to clarify the mechanisms by which new SNPs interfere with acetylation.

Suggested Reviewers: Tom Ottenhoff Ph.D  
Leiden University  
t.h.m.ottenhoff@lumc.nl

Brigitte Gicquel  
Institute Pasteur  
bgicquel@pasteur.fr

Tom Huizinga  
Leiden University  
t.w.j.huizinga@lumc.nl

Susan Dorman  
Johns Hopkins University  
dsusan1@jhmi.edu



**Rio de Janeiro, April, 15 2009**

Dear Editor Chief

Please find enclosed the manuscript entitled “Sequence analysis of *NAT2* gene of Brazilians: identification of undescribed Single Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure” from Raquel Lima de Figueiredo Teixeira et al. to be considered for publication as a research paper in the Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis.

In this manuscript we report the discovery of six new mutations within the *NAT2* coding region in Brazilians subjects and characterization of new *NAT2* alleles by cloning and re-sequencing followed by molecular modeling. According our *in silico* data, some of these SNPs and alleles are involved with the NAT2 low activity, contributing to the slow acetylation profile in these populations.

The authors declare that there are no conflicts of interest.

Yours sincerely,

---

Adalberto Rezende Santos, Ph.D.  
Laboratory of Molecular Biology Applied to Mycobacteria  
Oswaldo Cruz Institute – Fiocruz, Av. Brasil 4365, CEP: 21040-900  
Manguinhos, Rio de Janeiro, RJ, Brazil  
Phone: 55(21) 2598-4289 Fax: 55(21)2270-9997; E-mail: [adalbertorezende@yahoo.com.br](mailto:adalbertorezende@yahoo.com.br)

## *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*

### Conflict of Interest Policy

Article Title: Sequence analysis of *NAT2* gene of Brazilians: identification of undescribed Single Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure.

Author name: Adalberto Rezende Santos

### Declarations

*Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* requires that all authors sign a declaration of conflicting interests. If you have nothing to declare in any of these categories then this should be stated.

#### Conflict of Interest

A conflicting interest exists when professional judgement concerning a primary interest (such as patient's welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). It may arise for the authors when they have financial interest that may influence their interpretation of their results or those of others. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.

#### Please state any competing interests

Nothing to declare.

#### Funding Source

All sources of funding should also be acknowledged and you should declare any involvement of study sponsors in the study design; collection, analysis and interpretation of data; the writing of the manuscript; the decision to submit the manuscript for publication. If the study sponsors had no such involvement, this should be stated.

#### Please state any sources of funding for your research

Nothing to declare.

**Signature** (a scanned signature is acceptable, but each author must sign)

Adalberto Rezende Santos:

**Print name**

Adalberto Rezende Santos



## *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*

### Conflict of Interest Policy

Article Title: Sequence analysis of *NAT2* gene of Brazilians: identification of undescribed Single Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure

Author name: Raquel Lima de Figueiredo Teixeira

### Declarations

*Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* requires that all authors sign a declaration of conflicting interests. If you have nothing to declare in any of these categories then this should be stated.

#### Conflict of Interest

A conflicting interest exists when professional judgement concerning a primary interest (such as patient's welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). It may arise for the authors when they have financial interest that may influence their interpretation of their results or those of others. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.

#### Please state any competing interests

Nothing to declare.

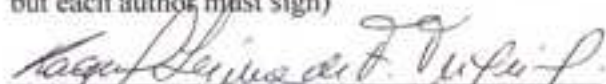
#### Funding Source

All sources of funding should also be acknowledged and you should declare any involvement of study sponsors in the study design; collection, analysis and interpretation of data; the writing of the manuscript; the decision to submit the manuscript for publication. If the study sponsors had no such involvement, this should be stated.

#### Please state any sources of funding for your research

Nothing to declare.

Signature (a scanned signature is acceptable, but each author must sign)



Print name

Raquel Lima de Figueiredo Teixeira

**Sequence analysis of *NAT2* gene of Brazilians: identification of undescribed Single Nucleotide Polymorphisms and molecular modeling of the N-acetyltransferase 2 protein structure**

Raquel L. F. Teixeira<sup>1</sup>, Floriano P. Silva Junior<sup>2</sup>, Aline R. Silveira<sup>2</sup>, Pedro H. Cabello<sup>1</sup>, Leila Mendonça-Lima<sup>3</sup>, Marcelo F Rabahi<sup>4</sup>, Afrânio L. Kritski<sup>5</sup>, Fernanda C. Q. Mello<sup>5</sup>, Philip N. Suffys<sup>6</sup>, Antonio B. Miranda<sup>3</sup> & Adalberto R. Santos<sup>6\*</sup>.

<sup>1</sup>Laboratory of Human Genetics, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

<sup>2</sup>Laboratory of Proteins and Peptides Biochemistry, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

<sup>3</sup>Laboratory for Functional Genomics and Bioinformatics, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

<sup>4</sup>Anuar Auad Infectious Disease Reference Hospital of the State Health Secretariat of Goiás State, Goiânia, Brazil

<sup>5</sup>Academic Tuberculosis Program / Medical School, Hospital Complex HUCFF-IDT, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>6</sup>Laboratory of Molecular Biology Applied to Mycobacteria, Oswaldo Cruz Institute, Fiocruz Rio de Janeiro, Brazil

**\*Corresponding author:** Laboratory of Molecular Biology Applied to Mycobacteria – Oswaldo Cruz Institute – Fiocruz, Av. Brasil 4365, CEP: 21040-900, Manguinhos, Rio de Janeiro, RJ, Brazil

**Phone:**55(21) 2598-4289 **Fax:** 55(21)2270-9997; **E-mail:** [adalbertorezende@yahoo.com.br](mailto:adalbertorezende@yahoo.com.br)

## Abstract

N-acetyltransferase 2 (NAT2) metabolizes a variety of xenobiotics that include many therapeutic drugs, chemicals and carcinogens. This enzyme is genetically variable in human populations and polymorphisms in the *NAT2* gene have been associated with drug toxicity and efficacy as well as cancer susceptibility. Here we have focused on the identification of *NAT2* variants in Brazilian individuals resident in two different regions by direct sequencing and in the characterization of new haplotypes after cloning and re-sequencing. Upon analysis of DNA samples from 404 individuals, six new mutations (c.29T>C, c.152G>T, c.203G>A, c.228C>T, c.458C>T and c.600A>G) and seven new *NAT2* alleles were identified with different frequencies in Rio de Janeiro and Goiás. Molecular modeling and structural analysis suggested that polymorphism p.Gly51Val may have an important effect on substrate recognition by NAT2. We also observed that mutation p.Cys68Tyr would impair acetylating activity due to the resulting geometric restrictions and incompatibility of the functional group in the Tyr side chain with the admitted chemical mechanism for catalysis by NATs. Moreover, other variants may lead to the presence of hydrophobic residues involved in protein aggregation and/or targeted degradation. From the new alleles, *NAT2*\*5N, \*5O, \*6H, \*6M, and \*12E may code for a slow activity and *NAT2*\*20 and \*12I, are functional alleles. Finally, it is possible that *NAT2*\*12F and *NAT2*\*12J do not interfere with the acetylation activity because they contain other mutations (c.622T>C and c.29T>C, in addition to c.803A>G) which do not affect the enzyme structure. Functional studies are needed to clarify the mechanisms by which new SNPs interfere with acetylation.

**Key words:** human NAT2, polymorphism, molecular modeling, N-acetylation, structure/function, Brazilian population.

## 1. Introduction

Arylamine N-acetyltransferases (NATs) constitute a ubiquitous enzyme class, which has been isolated from many organisms like bacteria, birds and mammals, and plays an important role for drug metabolism and in the bioactivation of carcinogenic arylamines. Humans express two functional NAT isozymes, NAT1 and NAT2, which share 81% amino acid sequence identity, but exhibit considerable differences in substrate specificity and tissue distribution [1].

Arylamine N-acetyltransferase II (NAT2) is a cytosolic phase II conjugation enzyme which catalyzes the transfer of an acetyl group from acetyl-CoA to the nitrogen or oxygen atom of different arylhydrazine and arylamine drugs [1, 2, 3, 4]. In addition, NAT2 catalyzes the metabolic activation of aromatic and heterocyclic amine carcinogens [1, 5].

Literature data have described the association of *NAT2* genetic variations and the acetylation status to the presence or absence of side effects during tuberculosis treatment [6, 7, 8; for a review see 9, 10] and also to susceptibility to bladder, colorectal and breast cancers [11, 12; for a review see 13]. In fact, three major genetically determined phenotypes are observed: fast, intermediate and slow acetylators, which are inherited as a codominant trait. Interindividual genetic variations in *NAT2* gene that cause differences in the activity of allelic isoenzymes or in the protein levels result in altered activation/deactivation of many compounds generating carcinogenic or toxic effects.

Most mutations reported to date are found within the 873 bp (including the stop codon) intronless coding region of *NAT2* gene, which is located in chromosome 8p22. Among the seven most common SNPs (single nucleotide polymorphisms), four result in amino acid changes leading to a significant decrease in acetylation capacity (SNPs c.191G>A, c.341T>C, c.590G>A and c.857G>A). The other three (c.282C>T, c.481C>T and c.803A>G) are silent

mutations and/or do not alter the phenotype [14, 15]. According to the consensus gene nomenclature of human *NAT2* ([www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html)), the *NAT2* alleles are characterized by the combination of up to four SNPs present in the coding region. Currently, there is a total of 32 SNPs identified in the *NAT2* coding region (25 in the nomenclature official site and the other seven in the Entrez SNP database) with 53 combinations (alleles) described.

A variation in the frequency of *NAT2* genotypes and phenotypes in different ethnic closed populations has been well described in different geographic regions around the world and include populations with high frequencies of slow-acetylator allele(s), such as European descendents and sub-Saharan Africans [16, 17, 18], and other populations that are characterized by higher frequencies of fast acetylator allele(s), including Japanese, Chinese and Amerindians [19, 20, 21]. According to Luca and co-workers, differences in eating habits and lifestyles may explain inter-population differences in *NAT2* variation [22].

Functional analysis and site directed mutagenesis studies have allowed the identification of the active site cysteine residue (corresponding to Cys68 in human NATs) as the only cysteine which was conserved in all of the other NATs of which sequence information was available. These studies have also shown that the same cysteine plays an essential role at the active site of human *NAT2* [23, 24]. Moreover, three-dimensional homology modeling studies of human *NAT2*, based on the *Salmonella typhimurium* NAT structure, have revealed a cysteine protease-like catalytic triad (Cys68-His107-Asp122) [25, 26].

Studies based on the homology of N-acetyltransferase enzyme crystal structures from different bacteria provided some understanding of its structural characteristics, including SNP locations and computational docking of substrates [25, 27, 28, 29, 30]. However, the alignment scores between bacterial and mammalian N-acetyltransferase protein sequences are

low (~30%) and important differences between the bacterial and mammalian N-acetyltransferase protein structures limited the utility of the mammalian N-acetyltransferase homology models [31].

Recently, the crystal structures of human NAT1 and NAT2 were finally resolved corroborating a number of predictions from previous modeling and experimental studies. The fold of human NAT1 and NAT2 closely resembles the overall structure of their prokaryotic orthologues, with three generally described domains: an 83 residues long N-terminal domain containing five  $\alpha$ -helices and a short  $\beta$ -strand; a second domain consisting of nine  $\beta$ -strands and two short helices forming a  $\beta$ -barrel; and a third  $\alpha/\beta$  lid domain with four  $\beta$ -strands and an  $\alpha$ -helix. The first two domains, which have been recognized to be involved in active-site pocket formation and AcCoA binding, are connected to the third domain through the  $\alpha$ -helical interdomain. The third domain has been implicated in acetyl-acceptor specificity [30, 32]. However, the crystal structures of human NATs also provided evidence for novel structural characteristics at the carboxyl terminus that could not be anticipated from the homology with the bacterial enzymes [32].

The crystal structures of human NAT1 and NAT2 have also shown that, although the point mutations are distributed throughout the entire primary structure, the majority of the SNPs are located on the protein surface and thus exposed to the cellular environment. These findings would be consistent with the hypothesis that the structural alterations induced by most amino acid substitutions basically lead to the ubiquitylation and proteasomal degradation of these proteins [33]. However, molecular modeling based on the recently solved human NATs structures aiming to give new insights about the role of SNPs on the acetylation phenotype has the potential to further expand these initial findings from structural analysis.

Genetic variations can alter the expression levels and/or function of a gene product and the haplotype structure is the main genetic determinant of phenotypic consequences.

Hence, we have been engaged on the identification of *NAT2* variants in Brazilian subjects by direct sequencing. We report here the discovery of six new mutations and seven new *NAT2* alleles by employing cloning and re-sequencing. Furthermore, molecular modeling and structural analysis were employed to identify polymorphisms with the potential to alter *NAT2* function. Such genetic variations should be useful markers for studying the correlation between genotypes and phenotypes of N-acetyltransferase 2.

## 2. Material and Methods

### Selection of subjects and sample collection

Four hundred and four consanguineously unrelated individuals from two different geographic regions of Brazil, Rio de Janeiro (RJ) and Goiás (GO), were enrolled in this study. They comprised 248 active TB patients and 50 asymptomatic close contacts of pulmonary TB cases from the Thoracic Institute/University Hospital Clementino Fraga Filho from the Federal University of Rio de Janeiro and 106 asymptomatic Health Care Workers from Anuar Auad Infectious Disease Reference Hospital of Goiás State, Brazil. This study was approved by the ethics committee of each enrolled institution.

After written consent, a volume of 3 mL of venous blood was collected from each of the volunteers and stored at  $-20^{\circ}\text{C}$ . Genomic DNA was isolated from 100  $\mu\text{L}$  of frozen whole blood using the FlexiGene DNA Kit (Qiagen Inc., USA), according to the manufacturer's specifications. After extraction, DNA samples were stored at  $-20^{\circ}\text{C}$ .

### *NAT2* genotyping and computational analysis

Genotyping of the *NAT2* coding region and computational analysis were performed as described previously [34]. In brief, a 1093 bp fragment, which contains the entire coding region of *NAT2*, was amplified by PCR and subsequently sequenced on both DNA strands using a set of primers for amplification (NAT2 EF 5' TTA GTC ACA CGA GGA AAT CAA A 3' and NAT2 ER 5' AAA TGC TGA CAT TTT TAT GGA TGA 3') and a second set of internal primers (NAT2 IF 5' ACC ATT GAC GGC AGG AAT TA 3' and NAT2 IR 5' TGG TCC AGG TAC CAG ATT CC 3') [34]. Sequencing was performed using ABI PRISM Big Dye Terminator v. 3.1 Kit (PE Applied BioSystems), according to the manufacturer's recommendations, on an ABI PRISM 3730 DNA Analyser (PE Applied BioSystems). The sequence analysis and SNPs identification were performed using SeqScape v.2.5 software



(Applied BioSystems <http://www.appliedbiosystems.com>) using GenBank Accession Number AY331807 ([http://www.ncbi.nlm.nih.gov/ GenBank](http://www.ncbi.nlm.nih.gov/GenBank)) as the reference sequence. Hardy-Weinberg equilibrium for each mutation was tested. All singletons and even new/rare mutation identified were confirmed by PCR re-amplification (three times) and re-sequencing.

### **Molecular determination of haplotypes by *NAT2* cloning and re-sequencing**

To resolve ambiguous haplotypes containing new or rare mutations, with 2 or more heterozygous SNPs, we used cloning methodology. Therefore, the 1093 bp PCR fragment of *NAT2* gene was cloned into pCR-BluntII-TOPO vector, to keep alleles separate, by using Zero Blunt TOPO PCR Cloning kit (Invitrogen) according to the manufacturer's recommendations. For these experiments, the final extension step of PCR cycle was abolished generating blunt ends in PCR fragments. Fifteen to twenty clones from each individual were isolated and plasmids were purified with Illustra plasmidPrep Mini Spin kit (GE Healthcare) and checked for the presence of insert by restriction analysis. Plasmids containing human *NAT2* insert were sequenced and the sequence data were analyzed with SeqScape software as described above.

### **Sequences and alignment**

The amino acid sequence of human *NAT2* (UniprotKB/Swiss-Prot accession number P11245) was employed as the query sequence to search the UniprotKB/Swiss-Prot databank [35] using the BLASTP program, as implemented in the WWW ExPASy server (<http://ca.expasy.org/>). Representative sequences from eukaryotic and prokaryotic organisms were selected and then aligned with the multiple sequence alignment tool T-COFFEE v.7.44 [36], using the program default parameters. The resulting alignment was manually corrected according to a structural alignment built by superposing in the Swiss-PdbViewer program v.3.7 [37] the crystal structures of NATs found in the Protein Data Bank (PDB) [38], including

human NAT2 chain A (PDB code 2FPR; [32]), human NAT1 (PDB code 2PQT; [32]) and *M. smegmatis* NAT chain A (PDB code 1W6F; [39]).

### **Molecular modeling and structural analysis of mutations**

For each SNP identified from the analysis of the human *NAT2* gene in Brazilian individuals, which resulted in amino acid substitution in the NAT2 enzyme, we have constructed a structural model of the mutated protein to evaluate the putative effects of each individual mutation. For this purpose, we have chosen chain A from the NAT2 crystal structure (PDB code 2FPR), since this protein was crystallized as a dimer [32]. The point mutations were introduced within the Swiss-Pdbviewer program by replacing the side chain of the mutated amino acid by browsing a rotamer library, which provides the most commonly observed orientations for the new side chain [37]. The most reasonable rotamers were chosen with the aid of a simple score based on the minimization of clashes and maximization of the number of hydrogen bonds between the mutated side chain atoms and other nearby atoms of the protein. We also favored rotamers that most preserved the space occupied by the original residue. The resulting models were submitted to an energy minimization protocol (to remove bad contacts and optimize favorable interactions) that consisted of 10 steps of the steepest-descent algorithm followed by up to 1000 steps of the conjugated gradient algorithm or until the energy gradient was inferior to 0.05 KJ/mol. In order to prevent over deviation from the crystal structure, harmonic constraints of 50/5000 for selected/non selected residues were employed, where the selected set was composed of any residue 6Å away from the mutated amino acid. The resulting optimized structures were inspected for any structural changes. Calculations of the electrostatic potentials on the molecular surfaces and the manual evaluation of interatomic contacts were done with built-in Swiss-Pdbviewer tools while an automated analysis of interatomic contacts was carried out within the LPC/CSU server [40].

### 3. Results and discussion

#### Identification of new mutations in *NAT2* coding region and new *NAT2* haplotypes

The sequencing of the 1093 bp amplicon containing the *NAT2* coding region obtained from 404 Brazilian individuals from two different regions, Rio de Janeiro and Goiás, allowed us to identify 17 polymorphisms already described in literature or data bases (Table 1) [34] and 6 new mutations (c.29T>C, c.152G>T, c.203G>A, c.228C>T, c.458C>T and c.600A>G). Among the SNPs already described in previous studies, we found four mutations (c.766A>G, c.622T>C, c.578C>T and c.472A>C) recently identified only in African populations [41, 42]. All beforementioned polymorphisms were observed at a frequency in both regions at Hardy-Weinberg equilibrium and their allele frequencies, found in the present study, as well as their presumed influence on the protein sequence are shown in Table 1. No significant differences were found in the frequencies of any SNP between individuals from Rio de Janeiro and Goiás (data not shown).

The new mutations in the *NAT2* coding region included transversion and transition events and only c.228C>T and c.600A>G are silent mutations. All new variants were singletons and some of them are present only in Rio de Janeiro (c. 29 T>C, c.152 G>T, c.203G>A, c.458C>T) and others only in Goiás (c.228C>T and c.600A>G) (Table 1). Upon comparison of the SNP profile in both regions we could observe that Goiás is more homogeneous than Rio de Janeiro showing a lower diversity of variants, probably as a consequence of the high influence of Amerindian population in this region, according with the history of Brazil's colonization ([www.ibge.gov.br/brasil500/index.html](http://www.ibge.gov.br/brasil500/index.html)).

To determine SNP combinations and then characterize new *NAT2* alleles at the DNA level, haplotypes ambiguity was resolved by isolating each *NAT2* allele through a cloning process and re-sequencing each chromosomal allele from recombinant plasmid DNA. We

were able to identify *NAT2\*6H*, *NAT2\*12E* and *NAT2\*12F* alleles, firstly identified in African populations [42], and seven new haplotypes with different profiles in the studied regions, as demonstrated in Figure 1.

### **Structural analysis of the SNPs identified in Brazilian individuals**

In order to have a global perspective of all known mutations affecting the coding region of the *NAT2* gene on the protein level, we have mapped the SNPs onto the 3D structure of the protein (Figure 2). Figure 2A shows the location of each SNP colored according to the occurrence in the Brazilian populations examined in this work. It is possible to see that all the mutations are well spread across the whole structure without showing a clear trend to cluster at a specific domain. From Figure 2B it is possible to appreciate the spatial relationship between mutations known to affect *NAT2* structure and/or activity, the new SNPs found in this work and the residues close to the 2-bromoacetanilide adduct with the catalytic Cys68 residue [modeled from the *NAT1* crystal structure (PDB code 2PQT)] [32]. Noteworthy, excluding p.Asp122Asn and possibly p.Gly286Glu and p.Arg64Gln/Trp, at least one half of the mutations with a documented phenotype (i.e. p.Leu137Phe, p.Ile114Thr and p.Arg197Gln) are found far away from the substrate binding site, indicating that they must produce their known unfavorable effects [10, 15, 43] through destabilization of the *NAT2* structure instead of having a direct effect over the interaction with the substrate. This trend is maintained among the new mutations identified in this work. The p.Cys68Tyr was the only mutation affecting a residue in direct contact with the substrate (Figure 2B).

### **Molecular modeling analysis of individual mutations**

We have employed a combination of sequence conservation analysis, inspection of molecular graphics and molecular modeling techniques to evaluate the possible effects of each SNP occurring in the coding region of the sequenced *NAT2* genes. The multiple sequence alignment used for this analysis can be found as a Supplementary Material on line.

Our analysis suggested that some of the new non-synonymous SNPs (c.458C>T, c.578C>T, c.683C>T, and c.838G>A) (p.Thr153Ile, p.Thr193Met, p.Pro228Leu and p.Val280Met, respectively) may lead to protein modifications by leaving hydrophobic residues highly exposed on the protein surface, thus facilitating protein aggregation and/or targeted degradation. Although the p.Thr193Met modification may lead to some destabilization of the  $\beta$ -barrel strands  $\beta$ 7 and  $\beta$ 8 in domain II, our analysis indicated that most of the identified SNPs should be innocuous to the function of the NAT2 enzyme, including c.29T>C, c.70T>A, c.403C>G, c.472A>C, c.609G>T, c.622T>C and c.766A>G (protein modifications p.Ile10Thr, p.Leu24Ile, p.Leu135Val, Ile158Leu, p.Glu203Asp, p.Tyr208His and p.Lys256Glu, respectively).

The p.Tyr208His mutation merits further attention as this mutation causes the bulkier Tyr side chain to be changed to the smaller His that can be well accommodated in the NAT2 structure (Figure 3). Our analysis showed that the imidazolyl group in His208 side chain could keep the hydrogen bond with one of the phosphate groups in CoA, although the hydrogen bonds with Ser12 and Thr214 were lost. Noteworthy, if protonation is assumed, the positively charged His208 would better accommodate the phosphate groups in CoA due to the Coulomb attraction of their negative charges. This view was supported by our model where it was predicted a shorter (stronger) hydrogen bond with the phosphate group in CoA (His208:NE2-CoA:O4A 2.6Å). Hence, we concluded that it is unlikely for the p.Tyr208His mutation to impair CoA binding by human NAT2. This conclusion is supported by the fact that His is also naturally found in the homologous position of chicken NAT2 sequence [44].

Next, we describe in detail the structural analysis of the individual mutations most likely affecting NAT2 function.

**c.152G>T (p.Gly51Val)**

Gly51 is conserved in many mammalian NATs but not in mycobacterial NAT sequences, where it is replaced by Ile in *M. smegmatis* or by Val in *M. tuberculosis* and *M. bovis*. It is positioned at the beginning of the  $\alpha 4$  helix in the human NAT2 structure and has 12.7% RSAS [solvent accessible surface area (SAS) of the residue in the folded protein relative to the SAS of the residue in vacuum]. Gly51 mainly contacts Ala54 and Ile55 in  $\alpha 4$  with 5.7% and 13% RCSA [contact surface area (CSA) with a particular partner relative to the total CSA of the residue with all contacting partners} respectively. During the modeling exercise, the introduction of the Val mutation had a difficult spatial accommodation (with only 3 possible rotamers) probably due the differences in side chain volumes ( $60.1\text{\AA}^3$  for Gly and  $140\text{\AA}^3$  for Val). The resulting model has doubled the contact surface with Glu49 (from 3.0% RCSA for Gly to 6.6% RCSA for Val). As this is a destabilizing contact (Val51:CG2-Glu49:OE1), the region between  $\alpha 2$  and  $\alpha 4$  can be dislocated, modifying the position of Phe37 side chain, which participates in substrate binding (Figure 4). Hence, this mutation may have an impact in substrate recognition by NAT2.

**c.203G>A (p.Cys68Tyr)**

NAT active site has been extensively studied and Cys68 was found to be essential for enzyme activity by site-directed mutagenesis and chemical modification [23, 24, 25, 45, 46]. Due to its importance to NAT function, Cys68 is conserved among prokaryotic and eukaryotic NATs. In human NAT2, this residue is located at the N-terminal end of the  $\alpha 5$  helix and is completely buried (0% RSAS) in a network of polar interactions (Figure 5A). Cys68 makes hydrogen bonds with Asn72 through the backbone (Cys68:O-Asn72:N  $3.0\text{\AA}$ ) and with His107 side chain through its sulfhydryl group (Cys68:SG-His107:ND1  $3.4\text{\AA}$ ). Hydrophobic residues make an extensive set of Van der Waals contacts: Phe37, Val71, Val106, Leu108 and Gly124 (0.9%, 2.4%, 5.3%, 4.8% and 3.1% RCSA, respectively). The

most important contacting partners are Asn72 (12.2% RCSA) and His107 (12.1% RCSA) side chains. Upon mutation to Tyr, two clashes were detected with residues Leu69 and Val106, but these could be readily eliminated after our energy minimization protocol. Regarding the impact of this SNP to the biochemical phenotype, although modeling of the p.Cys68Tyr mutation evidenced the availability of space to accommodate the larger Tyr side chain (Figure 5B), the phenol hydroxyl in Tyr would not fulfill the same role as the thiol group in Cys as a nucleophile. Hence, it's very unlikely that a NAT2 enzyme carrying the p.Cys68Tyr mutation can preserve its acetylation activity, at least not by the same catalytic mechanism, as we shall discuss further below.

There are two basic enzymatic mechanisms that are known to be used by acetyltransferases modifying a diverse set of substrates [47]. The first is through a direct transfer mechanism, as exemplified by the Gcn5 N-acetyl transferases (GNATs) superfamily, where a general base withdraws a proton from the substrate leaving a substrate atom more nucleophilic for the attack to the carboxyl moiety in AcetylCoA (more often, a His residue is activated by spatially close Glu or Asp to play the role of this base). The second is through a so-called ping-pong mechanism, as employed by the p300 and NAT families of acetyltransferases. Although structurally distinct, both families catalyze reactions that produce an intermediate, the acetyl-enzyme. In the ping-pong mechanism, acetyl-CoA binds and is attacked by a catalytic residue, resulting in the formation of an acetyl-enzyme covalent intermediate, with concomitant release of CoA. The second substrate binds to coordinate a nucleophilic attack on the acetyl-enzyme covalent intermediate. The acetyl group is then transferred and the modified substrate is released. Differently, for the direct transfer mechanism, the modified substrate is released followed by release of CoA.

The NAT family contains a motif, Cys-His-Asp, that is similar to the catalytic triad found in the cysteine proteases [48]. NATs and cysteine proteases use a similar mechanism

for their acetyltransferase and hydrolase activity, respectively, even though they use different substrates: NATs use acetyl-CoA and the amino group on arylamine, whereas cysteine proteases use a peptide and water. It was recently shown that acetylation of the hamster NAT2 is dependent on a thiolate-imidazolium ion pair, Cys-S-His-ImH<sup>+</sup>, and not a general acid-base catalysis [49], further highlighting the distinction from the direct transfer mechanism.

All of the knowledge briefly revised above strongly suggests that the p.Cys68Tyr mutation found in our study would prevent NAT2 from accelerating an acetylation reaction by nucleophilic catalysis in the known ping-pong mechanism. One hypothetical means by which the p.Cys68Tyr NAT2 mutant could retain its activity is the reaction mechanism changing to a direct transfer mechanism where His107 becomes a general base (still activated by Asp122) and Tyr68 plays a role as a general acid, facilitating CoASH departure. Such role for Tyr was described for the active site of Serotonin N-acetyltransferase [arylalkylamine N-acetyltransferase (AANAT)], a member of the GNAT superfamily [50, 51]. Nevertheless, a major obstacle for this hypothesis would remain - the lack of space in the NAT2 active site to bind both substrates simultaneously.

### **Molecular modeling analysis of selected alleles**

#### ***NAT2\*6H***

*NAT2\*6H* was the most frequently observed new allele among the studied population and was also the most prevalent allele in the group from Rio de Janeiro (Figure 1). The SNPs found in the coding region of this allele encode amino acid residues that are more than 40Å apart in the human NAT2 structure: p.Arg197Gln and p.Lys256Glu; thus it is very unlikely that one mutation could have direct influence over the other. The unfavorable effects to human NAT2 function produced by the p.Arg197Gln mutation are well known [10, 15]. Using a molecular modeling approach similar to ours, Walraven et al. [52] proposed that the slightly reduced thermostability frequently observed for the p.Arg197Gln variant, which leads to reduced



NAT2 activity and protein levels, are consistent with the loss of the relatively weak electrostatic interactions of Arg197 with Glu195 and Met105. There is no such effect nor any other phenotype change documented for the p.Lys256Glu mutation. Moreover, from our analysis we suggest that while the Lys256 variant reverses the charge in residue 256, its surface localization most likely prevents it from causing major alterations in the domain III stability or folding.

### ***NAT2\*5N***

*NAT2\*5N* was the second most frequent new allele for the totality of individuals examined as well as for the RJ group. For the group from Goiás, *NAT2\*5N* showed the same frequency of the *NAT2\*6H* allele. The *NAT2\*5N* allele codes for a triple mutant NAT2 variant: p.Ile114Leu, p.Ile158Thr and p.Lys268Arg. The p.Ile114Leu is a SNP frequently observed found in human populations and is associated with a slow acetylator phenotype [10, 15, 34]. Structural evaluations made by Walraven et al. [52] could offer no other reasoning for that functional observation other than an increase in protein aggregation and/or degradation due to the exposure of a hydrophobic residue on the protein surface. Although they are separated in space for more than 21Å, SNPs Ile114Thr and Ile158Leu are both located within the  $\beta$ -barrel domain and could sum up to have a synergic effect over human NAT2 structure and function. Nevertheless, residue 158 is found after a variable loop in NATs, at the N-terminal tip of the  $\beta$ 8 strand, and accordingly, our modeling for the mutation Ile158Leu suggested it would have little chance of affecting human NAT2 activity or stability. The third mutation found in the *NAT2\*5N* allele, p.Lys268Arg, was also analyzed by Walraven et al. [52] and these authors proposed that it could not affect the enzyme structure due to the complete lack of contacts with other protein residues, consistent with the functional data available so far [43]. Being at least 35Å away from either one of the sites of the other SNPs in the structure of this human NAT2 variant, we consider it is safe to presume that the

p.Lys268Arg mutation would have no direct influence on the phenotypes of the p.Ile114Thr and p.Ile158Leu SNPs.

### ***NAT2\*5O***

The other human NAT2 variant containing the p.Lys268Arg mutation is encoded by allele *NAT2\*5O*. The other two mutations found in this variant, p.Cys68Tyr and p.Ile114Thr, are both predicted to have a deleterious effect on NAT2 structure and function. While p.Ile114Thr was already characterized [10, 15], p.Cys68Tyr is described for the first time in this study. Due to its participation in the NAT2 catalytic triad and the different physical-chemical properties of Cys and Tyr, it is intuitive to assume that the p.Cys68Tyr mutation would have a major impact on NAT2 activity. The modeling studies reported above by us fully support this view. Because of the large distances separating one site of mutation from the other, it is unlikely that any of them would have a direct effect over the other. On the other hand, the concomitant expression of the p.Cys68Tyr and p.Ile114Thr mutations would most likely completely shut down the NAT2 variant encoded by this allele.

### ***NAT2\*12J, NAT2\*6M, NAT2\*12E and NAT2\*12F***

These four alleles were found exclusively in the Rio de Janeiro group, but in lower frequencies. All of them have in common the p.Arg197Gln SNP plus an additional mutation. From our molecular modeling analysis of the individual mutations, we propose that the alleles *NAT2\*6M* and *NAT2\*12E* may code for proteins with further reduced activity and/or stability. In the later allele, the side chains of residues 193 and 197, which are only 3.9Å apart, may produce a negative synergic effect on NAT2 activity upon mutation of both residues. As observed by Walraven et al. [52] the stabilization provided by the electrostatic interaction “sandwich” formed by the positively charged Arg197 positioned between the negatively charged Glu195 and the methionine sulfur’s electron pairs would be lost in the mutation for the neutral Gln. Our analysis suggests that the concomitant mutation of the polar Thr193 side

chain to the hydrophobic Met would further destabilize the Gln197 side chain, thus emphasizing the loss of thermostability produced by this mutation alone.

In conclusion, differential effects among several SNPs on the structure of NAT2 enzyme were observed suggesting that some SNPs, such as c.152G>T (p.Gly51Val) and c.203G>A (p.Cys68Tyr), may alter protein structure and/or binding to the catalytic site in the NAT2 protein. Molecular dynamics simulations are in progress in our group to further characterize such effects. However, functional studies are needed to clarify the dimension of the mechanisms by which these new identified SNPs interfere with acetylation. Improvements in our understanding of the relationships between polymorphic substitution at the NAT2 coding gene and the activity of the variant enzymes would be of great interest for design of new drugs. Moreover, associating SNPs with unfavorable therapeutic outcomes during treatment of human diseases has great potential for direct clinical application by providing new and more accurate genetic markers for diagnostic and prognostic purposes as well as novel therapeutic targets.

**Acknowledgments**

This work was supported by MCT/DECIT/MS and Faperj/Pronex Proc: E-26/170.0003/2008. Adalberto Rezende Santos, Afrânio Lineu Kritski, Fernanda C. Q. Mello, and Philip Noel Suffys are supported by CNPq Grant numbers 308786/2005-0, 35 04 77/1995-7, 306702/2007-0 and 312165/2006-4 respectively.

We thank the nursing service of the Anuar Auad Infectious Diseases Reference Hospital from the State of Goiás, and from the Hospital Tuberculosis Control Program, Thoracic Diseases Institute/Clementino Fraga Filho University Hospital for sample collection. We also thank the support from the PDTIS - FIOCRUZ Sequencing and Bioinformatics Platforms.

**Conflict of Interest statement**

The authors declare that there are no conflicts of interest.

## References

- [1] D.W. Hein, M.A. Doll, A.J. Fletland, M.A. Leff, S.J. Webb, G.H. Xiao, U.S. Devanaboyina, N.A. Nangju, Y. Feng. Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol. Biomarkers Prev.* **9** (2000) 29-42.
- [2] M.M. Reidenberg. The pharmacogenetics of antiarrhythmic drugs. *Ann. N. Acad. Sci.* **432** (1984) 69-74.
- [3] N. Sabbagh, E. Delaporte, D. Marez, J.M. Lo-Guidice, F. Piette, F. Broly. *NAT2* genotyping and efficacy of sulfasalazine in patients with chronic discoid lupus erythematosus. *Pharmacogenetics* **7** (1997) 131-135.
- [4] J.A. Trimbrell, J. Harland, V. Facchini. Polymorphic acetylation of hydralazine. *Clin. Pharmacol. Ther.* **28** (1980) 350-355.
- [5] D.W. Hein. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mutat. Res.* **506-507** (2002) 65-77.
- [6] M. Ohno, I. Yamaguchi, I. Yamamoto, T. Fukuda, S. Yolota, R. Maekura, M. Ito, Y. Yamamoto, T. Ogura, K. Maeda, K. Komuta, T. Igarashi, J. Azuma. Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity. *Int. J. Tuberc. Lung Dis.* **4** (2000) 256-261.
- [7] T. Kita, Y. Tanigawara, S. Chikazawa, H. Hatanaka, T. Sakaeda, F. Komada, S. Iwakawa, K. Okumura. N-acetyltransferase 2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients. *Biol. Pharm. Bull* **24** (2001) 544-549.
- [8] Y.S. Huang, H.D. Chern, W.J. Su, J.C. Wu, S.L. Lai, S.Y. Yang, F.Y. Chang, S.D. Lee. Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor antituberculosis drug-induced hepatitis. *Hepatology* **35** (2002) 883-889.

- [9] Y.S. Huang. Genetic polymorphisms of drug-metabolizing enzymes and the susceptibility to antituberculosis drug-induced liver injury. *Expert Opin. Drug Metab. Toxicol.* **3** (2007) 1-8.
- [10] P.D. Roy, M. Majumder, B. Roy. Pharmacogenomics of antiTB drugs-related hepatotoxicity. *Pharmacogenomics*, **9** (2008) 311-321.
- [11] A.C. Deitz, W. Zheng, M.A. Leff, M. Gross, W.Q. Wen, M.A. Doll, G.H. Xiao, A.R. Folsom, D.W. Hein. N-Acetyltransferase-2 genetic polymorphism, well-done meat intake, and breast cancer risk among postmenopausal women. *Cancer Epidemiol. Biomarkers Prev.* **9** (2000) 905-910.
- [12] K. Chen, Q.T. Jiang, H.Q. He. Relationship between metabolic enzyme polymorphism and colorectal cancer. *World J. Gastroenterol.* **11** (2005) 331-335.
- [13] D.W. Hein. N-acetyltransferase 2 genetic polymorphism: effects of carcinogen and haplotype on urinary bladder cancer risk. *Oncogene* **25** (2006) 1649–1658.
- [14] A.J. Fretland, M.A. Leff, M.A. Doll, D.W. Hein. Functional characterization of human N-acetyltransferase 2 (NAT2) single nucleotide polymorphisms. *Pharmacogenetics* **11** (2001) 207-215.
- [15] Y. Zang, M.A. Doll, S. Zhao, J.C. States, D.W. Hein. Functional characterization of single-nucleotide polymorphisms and haplotypes of human N-acetyltransferase 2. *Carcinogenesis* **28** (2007) 1665-1671.
- [16] C. Delomenie, L. Sica, D.M. Grant, R. Krishnamoorthy, J.M. Dupret. Genotyping of the polymorphic N-acetyltransferase (NAT2\*) gene locus in two native African populations. *Pharmacogenetics* **6** (1996) 177–185.
- [17] J.A.G. Agúndez. NAT2 genotyping: equilibrium between accuracy and feasibility in routine analyses. *J. Appl. Res.* **3** (2003) 1–6.

- [18] I. Cavaco, R. Reis, J.P. Gil, V. Ribeiro. *CYP3A4\*1B* and *NAT2\*14* alleles in a native African population. *Clin. Chem. Lab. Med.* **41** (2003) 606–609.
- [19] A. Sekine, S. Saito, A. Iida, Y. Mitsunobu, S. Higuchi, S. Harigae, Y. Nakamura. Identification of single-nucleotide polymorphisms (SNPs) of human N-acetyltransferase genes *NAT1*, *NAT2*, *AANAT*, *ARD1* and *LICAM* in the Japanese population *J. Hum. Genet.* **46** (2001) 314–319.
- [20] H.G. Xie, Z.H. Xu, D.S. Ou-Yang, Y. Shu, D.L. Yang, J.S. Wang, X.D. Yan, S.L. Huang, W. Wang, H.H. Zhou. Meta-analysis of phenotype and genotype of NAT2 deficiency in Chinese populations. *Pharmacogenetics* **7** (1997) 503–514.
- [21] L.F. Jorge-Nebert, M. Eichelbaum, E.U. Griese, T. Inaba, T.D. Arias. Analysis of six SNPs of NAT2 in Ngawbe and Embera Amerindians of Panama and determination of the Embera acetylation phenotype using caffeine. *Pharmacogenetics* **12** (2002) 39–48.
- [22] F. Luca, G. Bubba, M. Basile, R. Brdicka, E. Michalodimitrakis, O. Rickards, G. Vershubsky, L. Quintana-Murci, A.I. Kozlov, A. Novelletto. Multiple advantageous amino acid variants in the NAT2 gene in human populations. *Plos One* **3** (2008) 3136-3147.
- [23] J.M. Dupret and D.M. Grant. Site-directed mutagenesis of recombinant human arylamine N-acetyltransferase expressed in *Escherichia coli*. Evidence for direct involvement of Cys<sup>68</sup> in the catalytic mechanism of polymorphic human NAT2. *J. Biol. Chem.* **267** (1992) 7381-7385.
- [24] M. Watanabe, T. Sofuni, T. Nohmi. Involvement of Cys69 residue in the catalytic mechanism of N-hydroxyarylamine O-acetyltransferase of *Salmonella typhimurium*. Sequence similarity at the amino acid level suggests a common catalytic mechanism of acetyltransferase for *S. typhimurium* and higher organisms. *J Biol Chem* **267** (1992) 8429-8436.
- [25] J.C. Sinclair, J. Sandy, R. Delgoda, E. Sim, M.E.M. Noble. Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat. Struct. Biol.* **7** (2000) 560–564.

- [26] F. Rodrigues-Lima and J.M. Dupret. 3D model of human arylamine N-acetyltransferase 2: structure basis of the slow acetylator phenotype of the R64Q variant and analysis of the active-site loop. *Biochem. Biophys. Res. Commun.* **291** (2002) 116-123.
- [27] J. Sandy, A. Mushtaq, A. Kawamura, J. Sinclair, E. Sim, M. Noble. The structure of arylamine N-acetyltransferase from *Mycobacterium smegmatis* - an enzyme which inactivates the anti-tubercular drug, isoniazid. *J. Mol. Biol.* **318** (2002) 1071-1083.
- [28] S.J. Holton, J. Dairou, J. Sandy, F. Rodrigues-Lima, J.M. Dupret, M.E. Noble, E. Sim. Structure of *Mesorhizobium loti* arylamine N-acetyltransferase 1. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **61** (2005) 14-16.
- [29] I.M. Westwood, S.J. Holton, F. Rodrigues-Lima, J.M. Dupret, S. Bhakta, M.E.M. Noble, E. Sim. Expression, purification, characterization and structure of *Pseudomonas aeruginosa* arylamine N-acetyltransferase. *Biochem. J.* **385** (2005) 605-612.
- [30] I.M. Westwood, A. Kawamura, E. Fullam, A.J. Russell, S.G. Davies, E. Sim. Structure and mechanism of arylamine N-acetyltransferases. *Curr. Top. Med. Chem.* **6** (2006) 1641-1654.
- [31] J.M. Walraven, J.O. Trent, D.W. Hein. Computational and experimental analyses of mammalian arylamine N-acetyltransferase. Structure and function. *Drug Metab. Dispos.* **35** (2007) 1001-1007.
- [32] H. Wu, L. Dombrovsky, W. Tempel, F. Martin, P. Loppnau, G.H. Goodfellow, D.M. Grant, A.N. Plotnikov. Structural basis of substrate-binding specificity of human arylamine N-acetyltransferases. *J. Biol. Chem.* **282** (2007) 30189-30197.
- [33] F. Liu, N. Zhang, X. Zhou, P.E. Hanna, C.R. Wagner, D.M. Koepp, K.J. Walters. Arylamine N-acetyltransferase Aggregation and Constitutive Ubiquitylation. *J. Mol. Biol.* **361** (2006) 482-492.



- [34] R.L.F. Teixeira, A.B. Miranda, A.G.F. Pacheco, M.Q.P. Lopes, J. Fonseca-Costa, M. Rabahi, H.M. Melo, A.L. Kritski, F.C.Q. Mello, P.N. Suffys, A.R. Santos. Genetic profile of the arylamine N-Acetyltransferase 2 coding gene among individuals from two different regions of Brazil. *Mutat. Res.* **624** (2007) 31-40.
- [35] A. Bairoch, B. Boeckmann, S. Ferro, E. Gasteiger. Swiss-Prot: Juggling between evolution and stability. *Brief Bioinform* **5** (2004) 39-55.
- [36] C. Notredame, D.G. Higgins, J.T.C. Heringa. A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302** (2000) 205-217.
- [37] N. Guex and M.C. Peitsch. Swiss-PdbViewer: A Fast and Easy-to-use PDB Viewer for Macintosh and PC. *Protein Data Bank Quaterly Newsletter* **77** (1996) 7.
- [38] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. The Protein Data Bank. *Nucleic Acids Res.* **28** (2000) 235-242.
- [39] J. Sandy, S.H.E. Fullam, E. Sim, M. Noble. Binding of the anti-tubercular drug isoniazid to the arylamine N-acetyltransferase protein from *Mycobacterium smegmatis*. *Protein Sci.* **14** (2005) 775-782.
- [40] V. Sobolev, A. Sorokine, J. Prilusky, E.E. Abola, M. Edelman. Automated analysis of interatomic contacts in proteins. *Bioinformatics* **15** (1999) 327-332.
- [41] E. Patin, L.B. Barreiro, P.C. Sabeti, F. Austerlitz, F. Luca, A. Sajantila, D.M. Behar, O. Semino, A. Sakuntabhai, N. Guiso, B. Gicquel, K. McElreavey, R.M. Harding, E. Heyer, L. Quintana-Murci. Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. *Am. J. Hum. Genet.* **78** (2006) 423-436.
- [42] E. Patin, C. Harmant, K.K. Kidd, J. Kidd, A. Froment, S.Q. Mehdi, L. Sica, E. Heyer, L. Quintana-Murci. Sub-Saharan African coding sequence variation and haplotype diversity at the *NAT2* gene. *Hum. Mutat.* **27** (2006) 720-730.

- [43] Y. Zang, S. Zhao, M.A. Doll, J.C. States, D.W. Hein. Functional characterization of the A411T (L137F) and G364A (D122N) genetic polymorphisms in human N-acetyltransferase 2. *Pharmacogenet. Genomics* **17** (2007) 37-45.
- [44] M. Ohtomi, M. Sasaki, T. Deguchi. Two arylamine N-acetyltransferases from chicken pineal gland as identified by cDNA cloning. *Eur. J. Biochem.* **185** (1989) 253-261.
- [45] N.J. Butcher, A. Arulpragasam, R.F. Minchin. 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.* **279** (2004) 22131-22137.
- [46] H Wang, Z. Guo, G.M. Vath, C.R. Wagner, P.E. Hanna. Chemical modification of hamster arylamine N-acetyltransferase 2 with isozyme-selective and nonselective N-arylbromoacetamido reagents. *Protein J.* **23** (2004) 153-166.
- [47] S. Mukherjee, Y.H. Hao, K. Orth. A newly discovered post-translational modification - the acetylation of serine and threonine residues. *Trends Biochem. Sci.* **32** (2007) 210-216.
- [48] F. Rodrigues-Lima, C. Deloménie, G.H. Goodfellow, D.M. Grant, J.M. Dupret. Homology modeling 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.* **356** (2001) 327-334.
- [49] H. Wang, L. Liu, P.E. Hanna, C.R. Wagner. Catalytic mechanism of hamster arylamine N-acetyltransferase 2. *Biochemistry* **44** (2005) 11295-11306.
- [50] K.A. Scheibner, J. De Angelis, S.K. Burley, P.A. Cole. Investigation of the roles of catalytic residues in serotonin N-acetyltransferase. *J Biol Chem* **277** (2002) 18118-18126.
- [51] E. Wolf, J. De Angelis, E.M. Khalil, P.A. Cole, S.K. Burley. X-ray crystallographic studies of serotonin N-acetyltransferase catalysis and inhibition. *J. Mol. Biol.* **317** (2002) 215-224.

- [52] J.M. Walraven, Y. Zang, J.O. Trent, D.W. Hein. Structure/function evaluations of single nucleotide polymorphisms in human N-acetyltransferase 2. *Curr. Drug Metab.* **9** (2008) 471–486.
- [53] S. Fuselli, R.H. Gilman, S.J. Chanock, S.L. Bonatto, G. De Stefano, C.A. Evans, D. Labuda, D. Luiselli, F.M. Salzano, G. Soto, G. Vallejo, A. Sajantila, D. Pettener, E. Tarazona-Santos. Analysis of nucleotide diversity of *NAT2* coding region reveals homogeneity across Native American populations and high intra-population diversity. *Pharmacogenomics J.* **7** (2007) 144-152.
- [54] M.A. Leff, A.J. Fretland, M.A. Doll, D.W. Hein. Novel human N-acetyltransferase 2 alleles that differ in mechanism for slow acetylator phenotype. *J. Biol. Chem.* **274** (1999) 34519-34522.
- [55] D.W. Hein, A.J. Fretland, M.A. Doll. Effects of single nucleotide polymorphisms in human N-acetyltransferase 2 on metabolic activation (O-acetylation) of heterocyclic amine carcinogens. *Int. J. Cancer* **119** (2006) 1208-1211.
- [56] K. Shishikura, H. Hohjoh, K. Tokunaga. Novel allele containing a 190C>T nonsynonymous substitution in the N-acetyltransferase (NAT2) Gene. *Hum. Mutat.* **15** (2000) 581-584.
- [57] A. Sabbagh, A. Langaney, P. Darlu, N. Gérard, R. Krishnamoorthy, E.S. Poloni. Worldwide distribution of *NAT2* diversity: Implications for *NAT2* evolutionary history. *BMC Genetics* **9** (2008) 21-34.
- [58] E.G. Martin. Interethnic and intraethnic variability of NAT2 single nucleotide polymorphisms. *Curr. Drug Metab.* **9** (2008) 487-497.
- [59] N.M. Woolhouse, M.M. Qureshi, R.A. Bayoumi. A new mutation C759T in the polymorphic N-acetyltransferase (NAT2) gene. *Pharmacogenetics* **7** (1997) 83-84.

[60] A. Anitha and M. Banerjee. Arylamine N-acetyltransferase 2 polymorphism in the ethnic populations of South India. *Int. J. Mol. Med.* **11** (2003) 125-131.

[61] S. Boukouvala and E. Sim. Structural analysis of the genes for human arylamine N-acetyltransferases and characterization of alternative transcripts. *Basic Clin. Pharmacol. Toxicol.* **96** (2005) 343-351.

## Figure Legends

**Figure 1:** *NAT2* Haplotypes/alleles characterization containing new or rare mutations in Brazilian subjects. **A-** Human *NAT2* representative coding region. Figure was adapted from [61]. **B-** Haplotypes/alleles new or rare identified in this study. *NAT2\*6H*, *NAT2\*12E* and *NAT2\*12F* were first identified in African population [42]. The new haplotypes names are suggestions for the official site of *NAT2* nomenclature.

**Figure 2:** Mapping of the mutations on the coding region of the human *NAT2* gene on the three-dimensional structure of the protein. **A-** Spatial localizations of the residues undergoing mutations in the *NAT2* enzyme are represented by spheres centered at the  $\alpha$ -carbon coordinates in the human *NAT2* crystal structure (PDB code 2FPR). The protein folding is represented in light gray ribbons. The residues are labeled and colored according to their occurrence in the two different Brazilian populations (RJ population - yellow and co-occurrence in the RJ/GO populations - red). Mutations not identified in either one of the Brazilian populations are colored in gray. As a reference, the bound conformation of CoA, as observed in the human *NAT2* structure is shown as sticks colored according to the CPK scheme. **B-** All residues plotted in **A** are colored in red while the positions of residues possibly interacting with the arylamine substrate are marked by spheres colored in magenta (some of which are labeled in magenta for reference). The protein folding is represented by ribbons colored according to the domain structure of human *NAT2*: domain I ( $\alpha$ -helix bundle) – blue, domain II ( $\beta$ -barrel) - yellow, interdomain - gray and domain III ( $\alpha/\beta$ ) - green. Residues associated with SNPs that have a documented impact on *NAT2* stability and/or function are labeled in red. Residues associated with the new SNPs found within the Brazilian populations analyzed are detached in yellow boxes. Molecular graphics were prepared in the Swiss-Pdbviewer program [37].

**Figure 3:** Changes in the network of hydrogen bonds and in the electrostatic potential upon introduction of the Tyr208His mutation in human NAT2 structure. Molecular surfaces were colored according to the electrostatic potential (EP) calculated in the Swiss-Pdbviewer program [37]. The surface was rendered transparent to allow the visualization of the underlying residues (CPK colored sticks). Hydrogen bonds are represented as dashed green lines. The bound conformation of CoA as determined in the crystal structure of human NAT2 (PDB 2FPR) is also shown. Note that upon mutation of Tyr208 (**A**) to His (**B**), the imidazolyl group in the His208 side chain could keep the hydrogen bond with one of the phosphate groups in CoA, although the hydrogen bonds with Ser12 and Thr214 were lost. Observe also that in the His208 mutant the EP near the phosphate groups in CoA was slightly less negative.

**Figure 4:** Model for the Gly51Val mutation in human NAT2. In sticks (CPK colored) are shown Phe37, the acetanilide adduct with Cys68 (TYX) and all the residues 6Å away from Val51. The protein folding is represented by ribbons colored according to the domain structure of human NAT2: domain I ( $\alpha$ -helix bundle) – blue, domain II ( $\beta$ -barrel) - yellow, interdomain - gray and domain III ( $\alpha/\beta$ ) - green. The Cys68 adduct was modeled based on the crystal structure of the homologous human NAT1 (PDB 2PQT). The mutated Val51 is found in the N-terminal of the  $\alpha_4$  helix. Note that Phe37, which interacts with TYX is located at the tip of the loop connecting  $\alpha_4$  to  $\alpha_2$  and therefore may have its position affected by the p.Gly51Val mutation. Molecular graphics were prepared in the Swiss-Pdbviewer program [37].

**Figure 5:** Model for the Cys68Tyr mutation in human NAT2. The protein folding is represented by ribbons colored according to the domain structure of human NAT2: domain I ( $\alpha$ -helix bundle) – blue, domain II ( $\beta$ -barrel) - yellow, interdomain - gray and domain III ( $\alpha/\beta$ ) - green. In sticks are shown part of the bound CoA proximal to the catalytic site. **A-** Important residues contacting the Cys68 side chain are shown as sticks colored according to

the CPK scheme. The hydrogen bond between Cys68 and N72 backbones is represented as a dashed green line. **B-** The modeled Tyr68 side chain is shown along with the new conformation assumed by the Leu69 side chain (orange) needed to accommodate the bulkier phenol group in Tyr. Molecular graphics were prepared in the Swiss-Pdbviewer program [37].

Figure 1  
[Click here to download high resolution image](#)

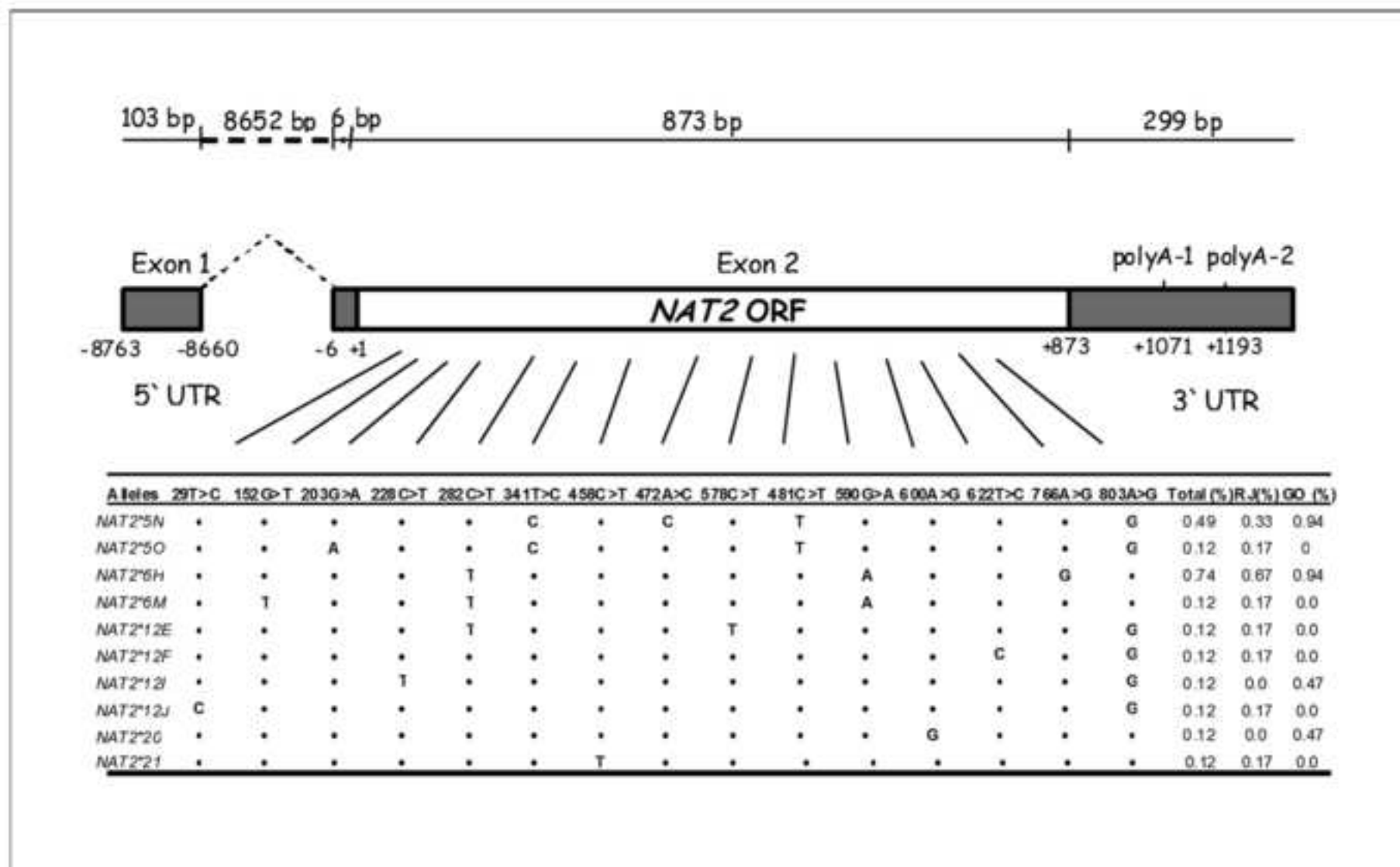




Figure 2A  
[Click here to download high resolution image](#)

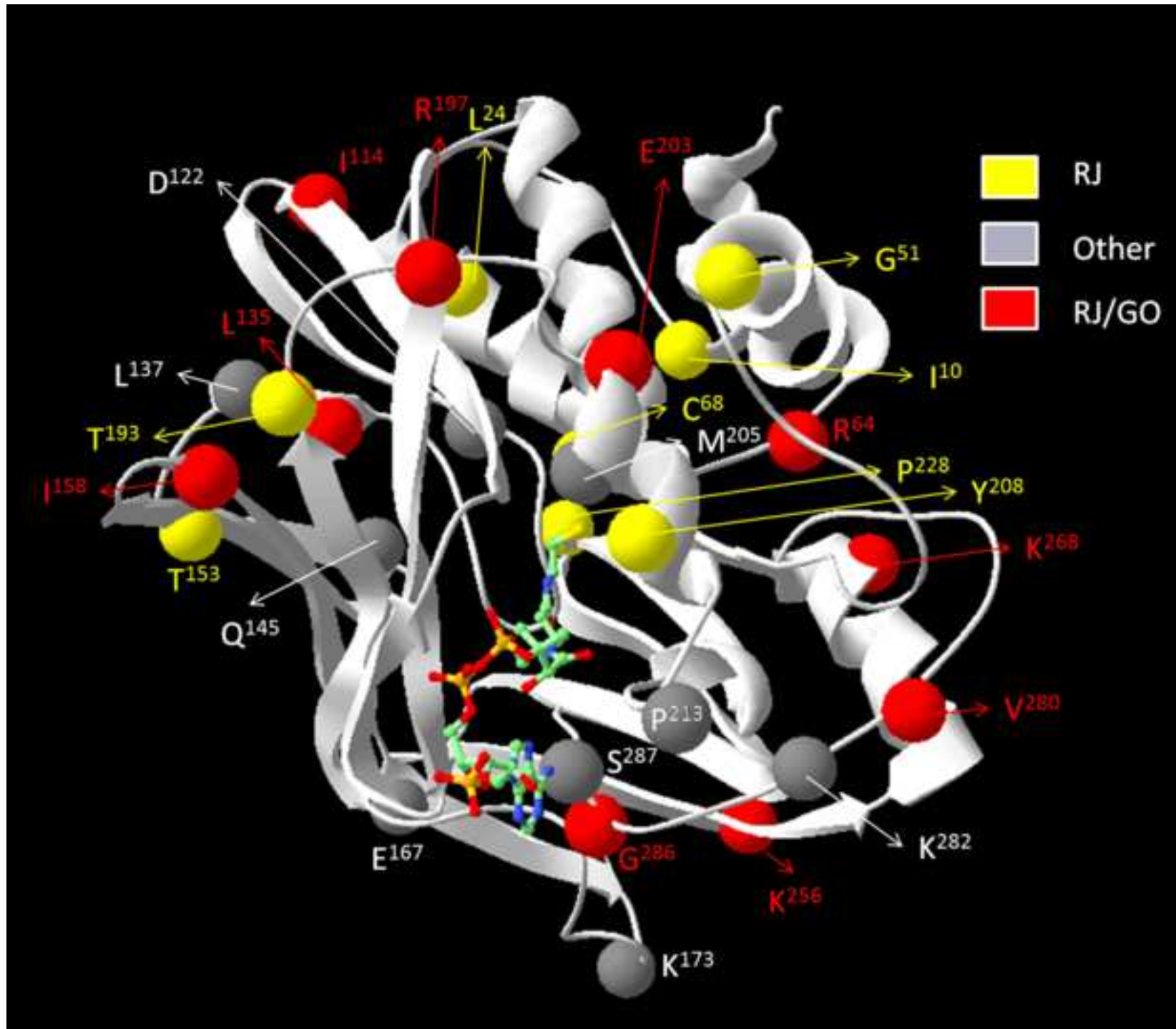


Figure 2B  
[Click here to download high resolution image](#)

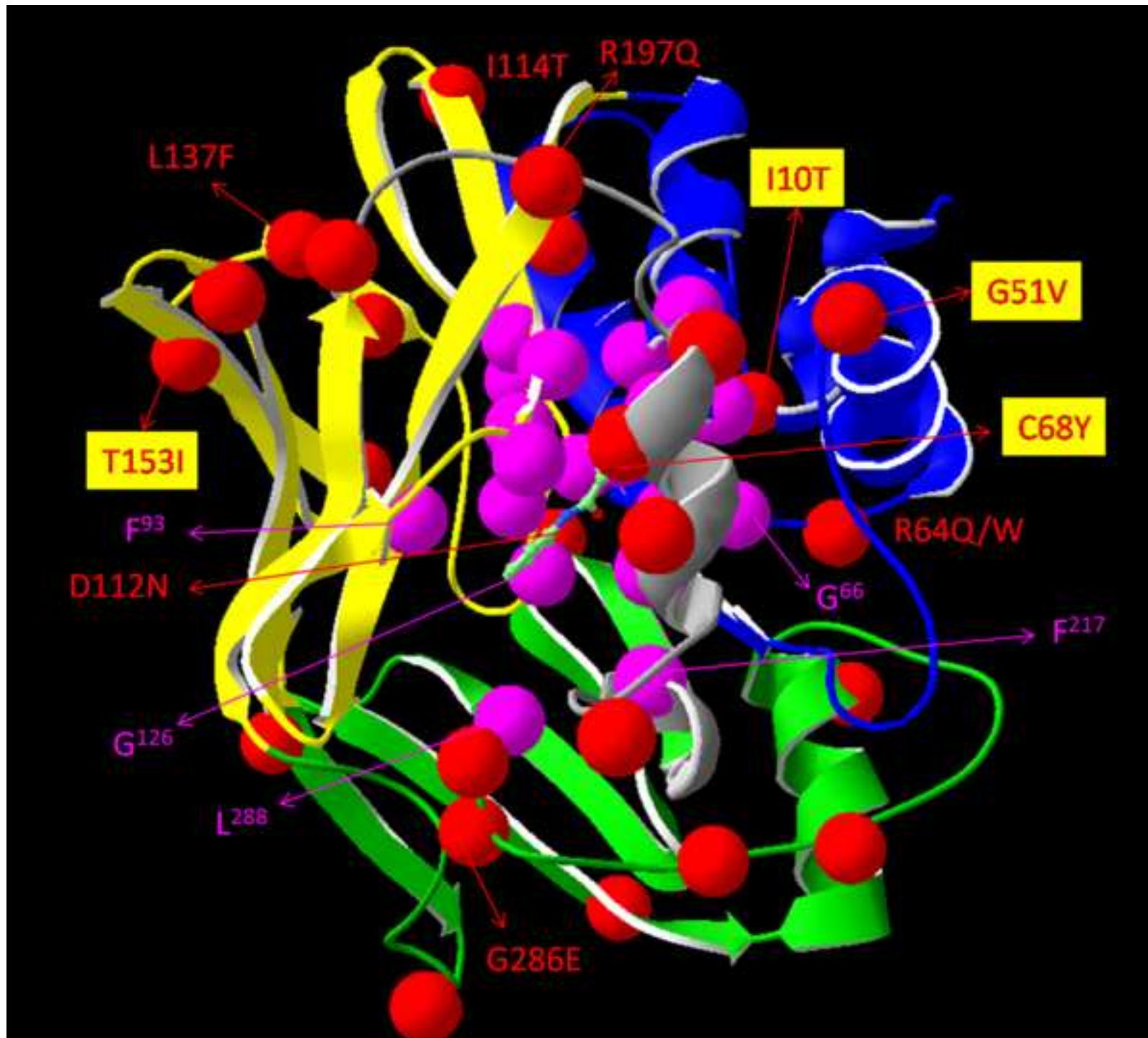


Figure 3  
[Click here to download high resolution image](#)

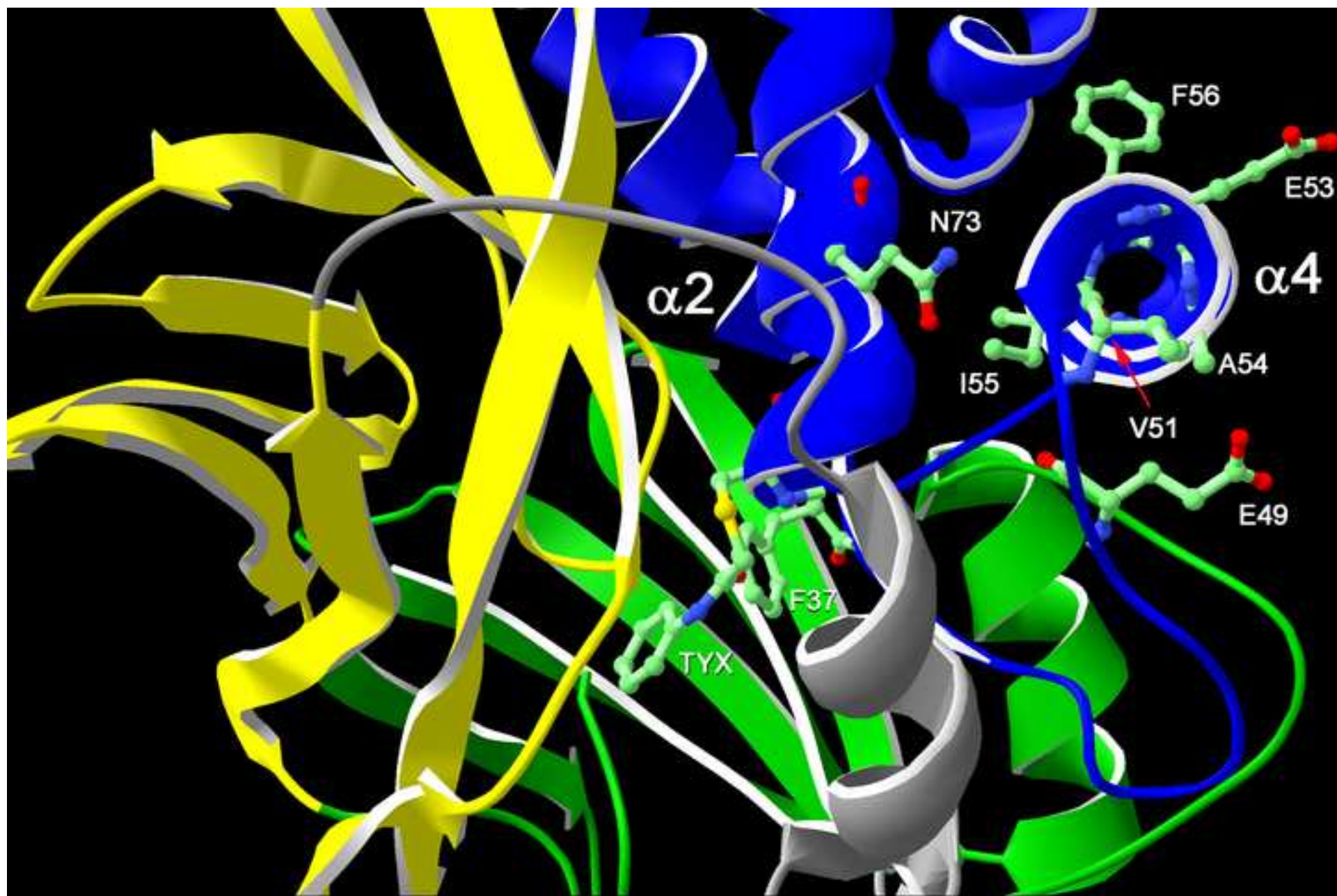
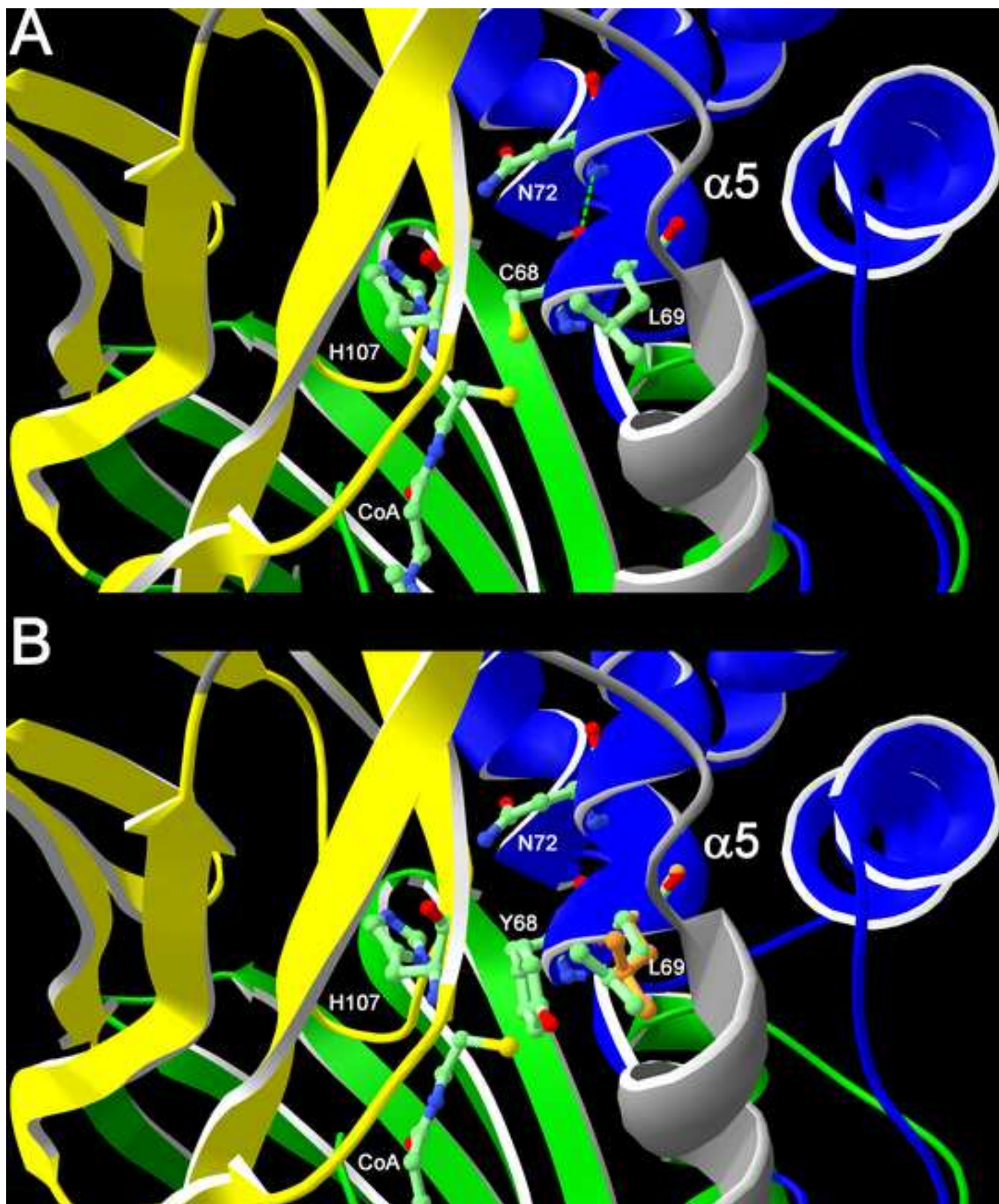
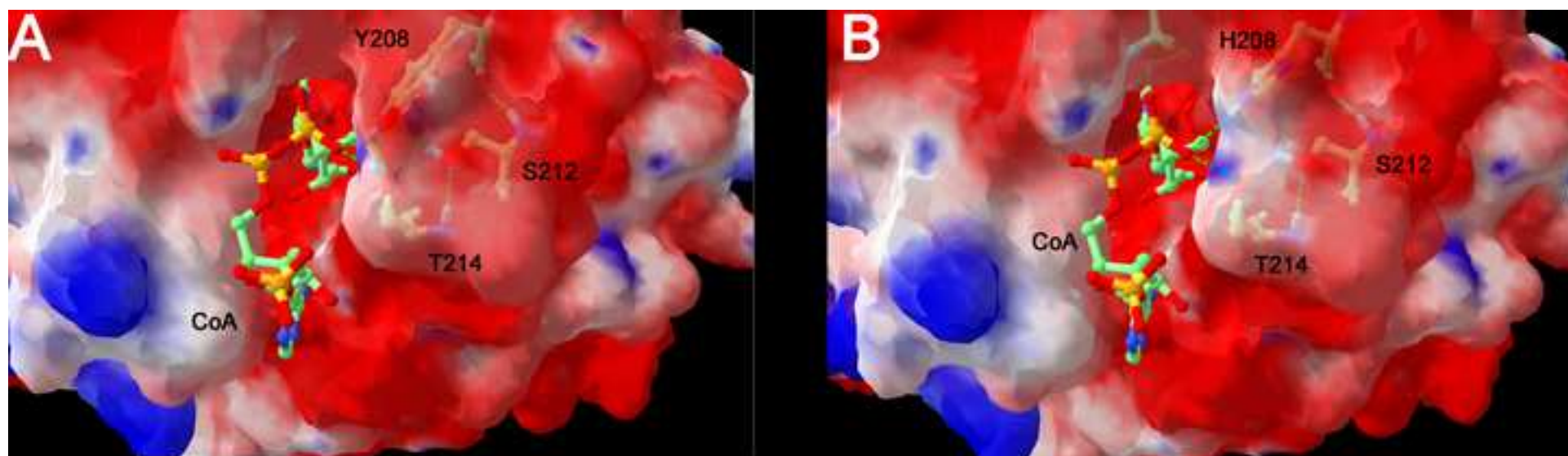


Figure 4  
[Click here to download high resolution image](#)



**Figure 5**  
[Click here to download high resolution image](#)



10 20 30 40 50 60 70 80  
ARY1\_HUMAN --MDIEAYLERIGYKKSERNKLDLETLTDILEHQIRAVPFENLN-IHCCDAMDGLGLBAIFDQVVRNRGGWCLQVNHLLYW 77  
ARY2\_HUMAN --MDIEAYFERIGYKNSRNKLDLETLTDILEHQIRAVPFENLN-MHCQAMELGLBAIFDHIIVRRNRGGWCLQVNHLLYW 77  
ARY1\_CHICK --MNLLEYFARTSYGGSLENDLETLTDIFOHIRAVPFENLN-IHCGETIILELHVYVNKIVHKRGGWCMENHLLGW 77  
ARY2\_CHICK --MDIKEYFARTSYGGSYEKPDLETLTEFOHHIQAVPFENLS-IHCGETIEMDLAATYDKIVRKRGGWCMENHLLGW 77  
ARY1\_MOUSE --MDIEAYFERIGYKNSVNLDLAFLTEVLOHQMRVAVPFENLN-MHCCEAMHLDLQDIFDHIIVRKRGGWCLQVNHLLYW 77  
ARY2\_MOUSE --MDIEAYFERIGYQSTRSKLDLKLTELELHQIRAVPFENLN-IHCESMELSLBAIFDQIVRKRGGWCLQVNHLLYW 77  
ARY1\_RABIT --MDIEAYQIRIGYKPNRNKLDLESLTDIFOHQIRVAVPFENLN-IHCESMELDLBAIFDQIVRNRGGWCLQVNHLLYW 77  
ARY2\_RABIT --MDIEAYQIRIGYKNRNKLDLESLTDIFOHQIRVAVPFENLS-IHCESMELDLBAIFDQIVRNRGGWCLQVNHLLYW 77  
ARY1\_MESAU --MDIEAYFERIGYNNPVYTLDLAFLTEVLOHQMRVAVPFENLN-MHCCEAMHLDLQDIFDHIIVRKRGGWCLQVNHLLYW 77  
ARY2\_MESAU --MDIEAYFERIGYQNSRNKLDLQTLTELELHQIRAVPFENLN-IHCESMELSLBAIFDQIVRKRGGWCLQVNHLLYW 77  
ARY1\_BOVIN --MDIEAYFERIGYKNSRDKLDLETLTDILEHQIRAVPFENLN-IHCCEAMHLDLQVIFDQIVRKRGGWCLQVNHLLYW 77  
ARY2\_MACMU --MDIEAYFERIGYKNSRNKLDLETLTDILEHQIRAVPFENLN-MHCCEAMHGLBTEFDHIIVRRNRGGWCLQVNHLLYW 77  
NAT\_MYCSM MAMDLDGGVLTTRIGLDG-RPRPDLGLLHAIVAAHNRSIIPFENLDPLLGPVADLSAALAKLVLRRRGGYCYEHNGLLGY 79  
NAT\_MYCTU MALDILTAYEDRINMYRG-ATDPTLDVLDLVTVHSRTIPFENLDPLLGVFVDDLSPOALADKLVLRRRGGYCFEHNGLMGY 79  
NAT\_MYCBO MALDILTAYEDRINMYRG-ATDPTLDVLDLVTVHSRTIPFENLDPLLGVFVDDLSPOALADKLVLRRRGGYCFEHNGLMGY 79

90 100 110 120 130 140 150 160  
ARY1\_HUMAN ALTTTIGFETMMLGGYVYSTPAK-KYSTGMIHLLQVITIDG--RNYIVDAGFC-RSYQMWPPELISGKDQPOVPCVFRIT 153  
ARY2\_HUMAN ALTTTIGFETMMLGGYFIPVPV-KYSTGMVHLLQVITIDG--RNYIVDAGSG-SSSQMWPELISGKDQPOVPCVFRIT 153  
ARY1\_CHICK VLKCLGYDTSFLGAYVFNPHEN-AMATMTTHLLVQVLEIG--KAYIVDAGFC-VSYQMWPPELISGKDQPOVPCVFRIT 153  
ARY2\_CHICK ALKKTGYNVYTLGAKVYIPEHD-AVADDIDHLLKVVYLDH--KSYIVDGGFC-MAYQLWQPELISGKDQPOVPCVFRIT 153  
ARY1\_MOUSE ALTKMGFETMMLGGYVYITPVS-KYSSBMVHLLVQVITISD--RKYIVDAGFC-GSYQMWPPELISGKDQPOVPAIFRIT 153  
ARY2\_MOUSE ALTKMGFETMMLGGYVITPPAN-KYSSGMHLLVQVITISG--KDYIVDAGFC-RSYQMWPPELISGKDQPOVPAIFRIT 153  
ARY1\_RABIT ALTTTIGFETMMLGGYVCGSHTD-KYSTGMIHLLVQVITING--RNYIVDAGFC-RSYQMWPPELISGKDQPOVPSIFRIR 153  
ARY2\_RABIT ALTTTIGFETMMLGGYVGSNND-KYSTGMIHLLVQVITING--RNYIVDAGFC-RSYQMWPPELISGKDQPOVPSIFRIR 153  
ARY1\_MESAU ALTMQGFETMMLGGYVYIVPVS-KYSSBMHLLVQVITISD--RNYIVDAAYC-GSYQMWPPELISGKDQPOVPAIFRIT 153  
ARY2\_MESAU ALTKMGFETMMLGGYVITPPAN-KYSSGMHLLVQVITISD--RNYIVDAGFC-RSYQMWPPELISGKDQPOVPAIFRIT 153  
ARY1\_BOVIN ALTTMIGFETMMLGGYVYNTFND-KYSSAMHLLKVTIDG--RDYIADAGFC-RSYQMWPPELISGKYQPOVPCVFRIT 153  
ARY2\_MACMU ALTTTIGFETMMLGGYVYIPAN-KYSTGMIHLLQVITIDG--RNYIADAGFC-SSSQMWPELISGKDQPOVPSIFRIT 153  
NAT\_MYCSM VLAELGFEVRRSAGRVMWRADAPLPAQTHNVLSVAVPCADGRYLVVDVGFQ-GQ-TLTSPLRLETCAVQPTTHEPYRLE 157  
NAT\_MYCTU VLAELGFEVRRSAGRVMWRADAPLPAQTHNVLSVAVPCADGRYLVVDVGFQ-GQ-TLTSPLRLETCAVQPTTHEPYRLE 157  
NAT\_MYCBO VLAELGYRVRFAARVWVKLADAPLPPQTHNVLSVAVPCADGRYLVVDVGFQ-GQ-TLTSPLRLETCAVQPTTHEPYRLE 157

170 180 190 200 210 220 230 240  
ARY1\_HUMAN EENGFWYLDQIRREQYIPNEEFHLSDDLLEDSYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_HUMAN BERGIWYLDQIRREQYITNKBELNSHLLPKKKHKQIYLFLEPRTEIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY1\_CHICK EKNAIYWLEKMRKQYIPNQNFSNDLLEKKDCRKYVMEFLPEPRTEIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_CHICK BENGWYLDQIRREQYIPNQNFSNDLLEKKNYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY1\_MOUSE BENGWYLDQIRREQYIPNEEFVNSDDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_MOUSE BENGWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY1\_RABIT BEGETWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_RABIT BEGETWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY1\_MESAU BENETWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_MESAU BENETWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY1\_BOVIN BDRGTWYLDQIRREQYIPNQBFINSDLLEKNKYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
ARY2\_MACMU BOKGIWYLDQIRREQYIPNQBFINSDLLEKNTYRKYIYSEFTLKPRTIEDFESMNTYLQTSPESSVETSKSFCSLQTPDGVHC 233  
NAT\_MYCSM RHGDDHTLAAQVRC-----EWQPLYTEFTTEPRPDIIEVGSWVSTHEGSHVETGLTVAVITDARYN 220  
NAT\_MYCTU DRVDGFVLQAMVRD-----TWQTLMEFTTQTRPQIDLKVASWMASTHEPASKVETGLTAAVITDARWN 220  
NAT\_MYCBO DRVDGFVLQAMVRD-----TWQTLMEFTTQTRPQIDLKVASWMASTHEPASKVETGLTAAVITDARWN 220

250 260 270 280 290 300  
ARY1\_HUMAN LVGFLLTHRRFNYKNDNDLVEFKTIS-BEELEKVLKNIENISLQR---KL---VPHKGRDFFT 289  
ARY2\_HUMAN LVGFLLTYRKFNYKNDNDLVEFKTIT-BEEVEVLKNIENISLGR---NL---VPPKGDGLT 289  
ARY1\_CHICK LIGWTLTETKYNKENDLVEFKITK-DEEVEKTLKDKENITLER---KL---VPINVKGFYT 289  
ARY2\_CHICK LVGKWLTEIKYNKNDNDLVEIRILA-DEEMKTLKDKENITLTK---KF---VPINTSRLSL 289  
ARY1\_MOUSE LVGSTFTSRFFSYKDDVDLVEFKYVN-BEELEDVLTAFGVSLEK---KF---VPHKGLVFT 289  
ARY2\_MOUSE LVGSTLTYRFFSYKDDVDLVEFKSIT-BEELEDVLTIEFVSLEK---KL---VPHKGRDFFT 289  
ARY1\_RABIT LVGLILTFRTYNKENNDLVEFKVLT-BEEVEVLKTIENISLGR---KL---VSKNGNLFFT 289  
ARY2\_RABIT LVGLILTFRTYNKENNDLVEFKVLT-BEEVEVLKTIENISLGR---KL---VSKNGNLFFT 289  
ARY1\_MESAU LIGSTIARRKFSYKENDLVEFKNVS-BEELEDVLTAFGVSLEK---KF---VPHKNGLSFS 289  
ARY2\_MESAU LVGSTLTYRFFSYKDDVDLVEFKSIT-BEELEDVLTIEFVSLEK---KL---VPHKGRDFFT 289  
ARY1\_BOVIN LVGFLLTYRFFNYKNDNDLVEFKTIN-EKELEENLNKNIENISLEK---KL---TPKHGDKFFT 289  
ARY2\_MACMU LVGFLLTYRFFNYKNDNDLVEFKTIN-EKELEENLNKNIENISLEK---KL---VPHKNGSFFT 289  
NAT\_MYCSM IRCRNLAVHR-AG-----ATGHIRFDSAAQVLDIVNREGIDLDGLAGR-DVQA-RVAEELLRDT-- 275  
NAT\_MYCTU ISCRDLAVHR-AG-----GTEKIRLADAAAQVDTLIEREGINVADIGERGALETREIDELLARQAP 279  
NAT\_MYCBO ISCRDLAVHR-AG-----GTEKIRLADAAAQVDTLIEREGINVADIGERGALETREIDELLARQAP 279

**Supplemental Figure S1:** Multiple sequence analysis of NAT sequences. Sequences are identified by their UniProtKB/Swiss-Prot entry names (UniProt Knowledgebase Release 14.7): ARY1\_HUMAN = Human NAT1 (UniProtKB accession number P18440; [62]); ARY2\_HUMAN = Human NAT2 (UniProtKB accession number P11245; [62]); ARY1\_CHICK = *Gallus gallus* (chicken) NAT1 (UniProtKB accession number P13913; [44]); ARY2\_CHICK = *G. gallus* (chicken) NAT2 (UniProtKB accession number P13914; [44]); ARY1\_MOUSE = *Mus musculus* (mouse) NAT1 (UniProtKB accession number P50294; [63]); ARY2\_MOUSE = *M. musculus* (mouse) NAT2 (UniProtKB accession number P50295; [63]); ARY1\_RABBIT = *Oryctolagus cuniculus* (Rabbit) NAT1 (UniProtKB accession number P18605; [64]); ARY2\_RABBIT = *O. cuniculus* (Rabbit) NAT2 (UniProtKB accession number P11246; [65]); ARY1\_MESAU = *Mesocricetus auratus* (Golden hamster) NAT1 (UniProtKB accession number P50292; [66]); ARY2\_MESAU = *M. auratus* (Golden hamster) NAT2 (UniProtKB accession number P50293; [67]); ARY1\_BOVIN = *Bos taurus* (Bovine) NAT1 (UniProtKB accession number Q1JPA6; [68]); ARY2\_MACMU = *Macaca mulatta* (Rhesus macaque) NAT2 (UniProtKB accession number Q7YRG5); NAT\_MYCSM = *Mycobacterium smegmatis* NAT (UniProtKB accession number O86309; [69]); NAT\_MYCTU = *M. tuberculosis* NAT (UniProtKB accession number P0A5L8; [69]); NAT\_MYCBO = *M. bovis* NAT (UniProtKB accession number P0A5L9; [70]). The alignment was shaded at the 70% conservation level according to the BLOSUM62 matrix [71] within the BioEdit software version 7.0.5.3 [72].

**References Supplementary Material:**

- [62] M. Blum, D.M. Grant, W. McBride, M. Heim, U.A. Meyer. Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression. *DNA Cell Biol.* **9** (1990) 193-203.
- [63] W.W. Weber, K.J. Martell, K.P. Vatsis. Molecular genetic basis of rapid and slow acetylation in mice. *Mol. Pharmacol.* **40** (1991) 218-227.
- [64] M. Blum, M. Heim, U.A. Meyer. Nucleotide sequence of rabbit NAT1 encoding monomorphic arylamine N-acetyltransferase. *Nucleic Acids Res.* **18** (1990) 5287-5287.
- [65] M. Blum, M. Heim, U.A. Meyer. Nucleotide sequence of rabbit NAT2 encoding polymorphic liver arylamine N-acetyltransferase (NAT). *Nucleic Acids Res.* **18** (1990) 5295-5295.
- [66] M. Abu-Zeid, K. Nagata, M. Miyata, S. Ozawa, M. Fukuhara, Y. Yamazoe. An arylamine acetyltransferase (AT-I) from Syrian golden hamster liver: cloning, complete nucleotide sequence, and expression in mammalian cells. *Mol. Carcinog.* **4** (1991) 81-88.
- [67] R.J. Ferguson, M.A. Doll, B.R. Baumstark, D.W. Hein. Polymorphic arylamine N-acetyltransferase encoding gene (NAT2) from homozygous rapid and slow acetylator congenic Syrian hamsters. *Gene* **140** (1994) 247-249.
- [68] G.P. Harhay, T.S. Sonstegard, J.W. Keele, M.P. Heaton, M.L. Clawson, W.M. Snelling, R.T. Wiedmann, C.P. Van Tassell, T.P.L. Smith. Characterization of 954 bovine full-CDS cDNA sequences. *BMC Genomics* **6** (2005) 166-166.
- [69] M.A. Payton, R. Auty, R.T. Delgoda, M. Everitt, E. Sim. Cloning and characterization of arylamine N-acetyltransferase genes from *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: increased expression results in isoniazid resistance. *J. Bacteriol.* **181** (1999) 1343-1347.



- [70] T. Garnier, K. Eiglmeier, J.C. Camus, N. Medina, H. Mansoor, M. Pryor, S. Duthoy, S. Grondin, C. Lacroix, C. Monsempe, S. Simon, B. Harris, R. Atkin, J. Doggett, R. Mayes, L. Keating, P.R. Wheeler, J. Parkhill, B.G. Barrell, S.T. Cole, S.V. Gordon R.G. Hewinson. The complete genome sequence of *Mycobacterium bovis*. Proc. Natl. Acad. Sci. USA **100** (2003) 7877-7882.
- [71] JG Henikoff. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA **89**: (1992) 10915–10919.
- [72] T.A. Hall. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. **41** (1999) 95-98.

Table 1: All mutations identified in *NAT2* coding region and their frequencies among 404 Brazilian subjects from Rio de Janeiro and Goiás.

Mutation	References	rs or ss identifiers**	Effect	aa change	Frequency (%)		
					Total <sup>a</sup>	RJ <sup>b</sup>	GO <sup>c</sup>
<b>c.29T&gt;C*</b>	<b>This study</b>	<b>ss102664284</b>	<b>missense</b>	<b>p.Ile10Thr</b>	<b>0.12</b>	<b>0.17</b>	<b>0.0</b>
c.34T>C	22	-	missense	p.Tyr12His	-	-	-
c.70T>A	34, 41, 53	rs45477599	missense	p.Leu24Ile	0.12	0.17	0.0
c.111T>C	54, 55	-	silent	-	-	-	-
<b>c.152G&gt;T*</b>	<b>This study</b>	<b>ss102664285</b>	<b>missense</b>	<b>p.Gly51Val</b>	<b>0.12</b>	<b>0.17</b>	<b>0.0</b>
c.190C>T	52, 56	rs1805158	missense	p.Arg64Trp	-	-	-
c.191G>A	34, 43, 52	rs1801279	missense	p.Arg64Gln	4.08	4.7	2.36
<b>c.203G&gt;A*</b>	<b>This study</b>	<b>ss102664286</b>	<b>missense</b>	<b>p.Cys68Tyr</b>	<b>0.12</b>	<b>0.17</b>	<b>0.0</b>
<b>c.228C&gt;T*</b>	<b>This study</b>	<b>ss102664287</b>	<b>silent</b>	<b>-</b>	<b>0.12</b>	<b>0.0</b>	<b>0.47</b>
c.282C>T	14, 15, 34	rs1041983	silent	-	36.88	37.92	33.96
c.341T>C	15, 34, 52	rs1801280	missense	p.Ile114Thr	37.87	39.6	33.02
c.345C>T	34, 57	rs45532639	silent	-	0.25	0.0	0.94
c.364G>A	15, 43, 52	rs4986996	missense	p.Asp122Asn	-	-	-
c.403C>G	34, 42, 57	rs12720065	missense	p.Leu135Val	0.62	0.67	0.47
c.411A>T	15, 43, 52	rs4986997	missense	p.Leu137Phe	-	-	-
c.434A>C	14, 52, 55	-	missense	p.Gln145Pro	-	-	-
<b>c.458C&gt;T*</b>	<b>This study</b>	<b>ss102664288</b>	<b>missense</b>	<b>p.Thr153Ile</b>	<b>0.12</b>	<b>0.17</b>	<b>0.0</b>
c.472A>C	42 and this study	-	missense	p.Ile158Leu	0.62	0.50	0.94
c.481C>T	14, 15, 34	rs1799929	silent	-	35.15	36.41	31.60
c.499G>A	15, 52, 55	-	missense	p.Glu167Lys	-	-	-
c.518A>G	42	-	missense	p.Lys173Arg	-	-	-
c.578C>T	41; 53 and this study	-	missense	p.Thr193Met	0.12	0.17	0.0
c.590G>A	14, 34, 52	rs1799930	missense	p.Arg197Gln	26.61	26.01	28.30
<b>c.600A&gt;G*</b>	<b>This study</b>	<b>ss102664289</b>	<b>silent</b>	<b>-</b>	<b>0.12</b>	<b>0.0</b>	<b>0.47</b>
c.607G>C	53	-	missense	p.Glu203Gln	-	-	-
c.609G>T	42, 57	rs45618543	missense	p.Glu203Asp	0.37	0.33	0.47
c.613A>T	58	rs45607939	missense	p.Met205Leu	-	-	-
c.622T>C	41, 53 and this study	rs56387565	missense	p.Tyr208His	0.12	0.17	0.0
c.638C>T	57	-	missense	p.Pro213Leu	-	-	-
c.683C>T	34, 42	rs45518335	missense	p.Pro228Leu	0.25	0.33	0.0
c.759C>T	55, 59	rs56011192	silent	-	-	-	-
c.766A>G	42 and this study	rs55700793	missense	p.Lys256Glu	0.74	0.67	0.94
c.803A>G	15, 34, 52	rs1208	missense	p.Lys268Arg	39.85	41.11	36.32
c.838G>A	34, 42, 57	rs56393504	missense	p.Val280Met	0.49	0.50	0.47
c.845A>C	14, 55; 52	rs56054745	missense	p.Lys282Thr	-	-	-
c.857G>A	15, 34, 52	rs1799931	missense	p.Gly286Glu	4.45	4.53	4.24
c.859T>C	13	-	missense	p.Ser287Pro	-	-	-
c.859delT	60	-	Frameshift	p.Ser287fs	-	-	-

\* New mutations identified in *NAT2* coding region. \*\* Reference cluster ID numbers (rs) or NCBI Assay ID numbers (ss) in dbSNP Entrez.

<sup>a</sup> Frequency refers to 808 alleles. <sup>b</sup> Frequency refers to 596 alleles. <sup>c</sup> Frequency refers to 212 alleles.



## CAPÍTULO IV

Essa etapa incluiu o estudo de associação, do tipo caso-controle, entre os polimorfismos genéticos presentes nos genes que codificam para as enzimas metabolizadoras de drogas NAT2, CYP2E1, GSTM1 e GSTT1, com a ocorrência de reações adversas em pacientes com TB em tratamento com esquemas contendo isoniazida. Esses resultados encontram-se reunidos no manuscrito “Genetic polymorphisms of NAT2, CYP2E1, GSTM1 and GSTT1 enzymes and the occurrence of antituberculosis drug-induced hepatotoxicity in Brazilian TB patients” **submetido** à revista internacional *European Journal of Human Genetics* em abril de 2009.

**Genetic polymorphisms of NAT2, CYP2E1, GSTs enzymes and the occurrence of  
antituberculosis drug-induced hepatotoxicity in Brazilian TB patients**

**Running title:** Drug-metabolizing enzymes and hepatotoxicity

Raquel L. F. Teixeira<sup>1</sup>, Renata G. Morato<sup>1</sup>, Pedro H. Cabello<sup>1</sup>, Ligia M. K. Muniz<sup>2</sup>, Adriana S. R. Moreira<sup>2</sup>, Afrânio L. Kritski<sup>2</sup>, Fernanda C. Q. Mello<sup>2</sup>, Philip N. Suffys<sup>3</sup>, Antonio B. Miranda<sup>4</sup> & Adalberto R. Santos<sup>3\*</sup>.

<sup>1</sup>Laboratory of Human Genetics, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

<sup>2</sup>Medical School -Hospital Complex HUCFF-IDT, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

<sup>3</sup>Laboratory of Molecular Biology Applied to Mycobacteria, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

<sup>4</sup>Laboratory of Functional Genomic and Bioinformatics, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil

**\*Corresponding author:** Laboratory of Molecular Biology Applied to Mycobacteria – Oswaldo Cruz Institute – Fiocruz, Av. Brasil 4365, CEP: 21040-900, Manguinhos, Rio de Janeiro, RJ, Brazil. **Phone:**55 (21) 2598-4289 **Fax:** 55 (21)2270-9997

**E-mail:** [adalbertorezende@yahoo.com.br](mailto:adalbertorezende@yahoo.com.br)

**Abstract**

The problem of liver adverse reactions induced by antituberculosis drugs is a relevant subject of investigation. Isoniazid (INH), one of the most important drugs used in antituberculosis treatment, is also the major drug involved in hepatotoxicity. Differences in isoniazid-induced toxicity have been attributed to genetic variability in several loci such as *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1*, coding for drug-metabolizing enzymes. Considering the differences in the *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1* alleles and distribution among interethnic populations, our goal was to evaluate the contribution of their polymorphisms in the susceptibility to antituberculosis drug-induced hepatotoxicity in Brazilian individuals. A total of 167 TB patients treated with isoniazid-containing antituberculosis schemes were genotyped for *NAT2*, *CYP2E1*, *GSTM1* and *GSTM1* polymorphisms. Genotype and allele frequencies were compared between TB patients with (cases) or without liver adverse events using multivariate logistic regression analyses. Among cases, 37 patients presented hepatotoxicity and 27 had drug-induced hepatitis. Statistical analyses revealed no significant association between *CYP2E1*, *GSTM1* and *GSTT1* genotypes and liver side effects occurrence. However, slow acetylators had a higher incidence of hepatitis than intermediate/fast acetylators [22% (18/82) versus 9.8% (6/61); OR, 2.86; CI 95%, 1.06-7.68;  $p = 0.04$ ]. Logistic regression showed that slow acetylation status was the only independent risk factor for both antituberculosis drug-induced hepatotoxicity (OR, 2.62; 95% IC, 1.75-3.49;  $p = 0.03$ ) and hepatitis (OR, 3.59; 95% IC, 2.53-4.64;  $p=0.02$ ). Therefore, slow acetylation status is a significant risk factor for hepatotoxicity and/or hepatitis occurrence during antituberculosis treatment with isoniazid-containing schemes in Brazilian individuals.

**Keywords:** Drug-metabolizing enzymes, hepatotoxicity, tuberculosis, isoniazid, pharmacogenetics, Brazilian subjects

## Introduction

Tuberculosis (TB) is one of the major causes of illness and death worldwide, especially in Asia and Africa. Globally, 9.2 million new cases and 1.7 million deaths from TB occurred in 2006, of which 0.7 million cases and 0.2 million deaths were in HIV-positive individuals.<sup>1</sup> Brazil is the 16<sup>th</sup> among the 22 countries with the higher burden of tuberculosis, with 50 million infected subjects, 110 000 new cases and 6 000 annual deaths.<sup>2,3</sup>

According to the guidelines for TB control, treating patients with anti-TB drugs is one of the most effective intervention and schemes containing isoniazid (INH), rifampicin (RMP), ethambutol (EMB) and pyrazinamide (PZA) are used as first line therapy for tuberculosis<sup>1,2</sup> Despite the availability of effective chemotherapeutic agents to treat TB, side effects such as hepatic disorder and peripheral neuritis are frequent, limiting the clinical use of the drugs and contributing to treatment relapse and appearance of multidrug-resistant *Mycobacterium tuberculosis* strains.<sup>4,5</sup>

The most frequent and serious side effect of these drugs is hepatotoxicity.<sup>4,6</sup> The liver has a central role in drug metabolism and detoxification, and is consequently vulnerable to injury. Among the first line anti-TB drugs, INH is the main responsible for the occurrence of drug-induced hepatic adverse effects and the metabolic intermediates of INH seem to be the cause of hepatotoxicity.<sup>7-9</sup> In liver, INH is metabolized to acetylisoniazid via N-acetyltransferase 2 (NAT2), followed by hydrolysis to acetylhydrazine. Acetylhydrazine is oxidized by cytochrome P4502E1 (CYP2E1) to form hepatotoxic intermediates.<sup>8,10</sup> These reactive metabolites can destroy hepatocytes either by interfering with cell homeostasis or by triggering immunologic reactions in which reactive metabolites, bound to hepatocyte plasma proteins, may act as haptens.<sup>11</sup> The other metabolic pathway to generate toxic metabolites is the direct hydrolysis of INH to hydrazine, a potent hepatotoxin. N-acetyltransferase 2 is also responsible to convert acetylhydrazine to diacetylhydrazine, a nontoxic component.<sup>7-9</sup> Glutathione S-transferase (GST), an important phase II detoxification enzyme, is supposed to play a protective role as an intracellular free radical scavenger conjugating glutathione with toxic metabolites that

are generated from CYP2E1.<sup>12</sup> Sulphydryl (SH) conjugation facilitates the elimination of metabolites from the body and reduces the toxic effect.<sup>13</sup>

The common risk factors for hepatotoxicity include advanced age, sex, malnutrition, complications of diseases, alcohol abuse, HIV infection, viral hepatitis B and C infections.<sup>6, 14, 15</sup> Recent studies have demonstrated that genetic polymorphisms in drug-metabolizing enzymes, which influence their activation capacity and the generation of metabolites in the liver, might predispose an individual to hepatic adverse reactions.<sup>16-24</sup>

*NAT2* is a polymorphic gene and mutations in its coding region can alter its enzymatic activity.<sup>25, 26</sup> In fact, three major genetically determined phenotypes are observed: fast, intermediate and slow acetylators, which are inherited as a codominant trait.<sup>27-29</sup> However, some groups consider the intermediate and fast acetylators as one and classify genotypes as: slow and fast acetylators.<sup>30-32</sup> The wild-type *NAT2*\*4 allele and 52 other variants of *NAT2* gene were identified and classified in human populations depending on the haplotype determined by the combination of up to four out of the 25 different single nucleotide polymorphisms (SNPs) present throughout the *NAT2* coding region ([www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html)). Several other rare SNPs were described in other populations and reported (<http://www.ncbi.nlm.nih.gov/GenBank>), but they have not been taken into account for allele determination. Many studies have demonstrated that people with low *NAT2* activity have a higher risk of developing hepatic disorders than those with high *NAT2* activity.<sup>16, 17, 19, 20</sup> However, other authors observed that fast acetylation would be a factor risk for antituberculosis drug-induced hepatotoxicity and the controversy still remains.<sup>33-36</sup>

The activity of CYP2E1 is also modulated by polymorphisms at several sites on its gene and the higher activity of this enzyme may increase the synthesis of hepatotoxins leading to hepatotoxicity. Two polymorphisms upstream of the *CYP2E1* transcriptional start site are detectable by *Pst* I and *Rsa* I restriction enzymes and appear to be in complete linkage disequilibrium. Three genotypes of *CYP2E1* are classified as c1/c1, c1/c2 and c2/c2 by

Restriction Fragment Length Polymorphism (RFLP) using *RsaI* as the restriction enzyme. The wild-type allele is c1 and the mutant allele is c2. According to the recent recommended nomenclature for *CYP2E1*, \*1A is equivalent to c1, and \*5 is equivalent to c2 ([www.cypalleles.ki.se/cyp2e1.htm](http://www.cypalleles.ki.se/cyp2e1.htm)). The presence of *CYP2E1* \*1A/\*1A as a genetic marker of isoniazid-induced liver injury has been evaluated in few studies with discordant results.<sup>21, 36, 37</sup>

The absence of GST activity due to homozygous null mutations at *GSTM1* and *GSTT1* loci may also modulate susceptibility to drug-induced hepatotoxicity, but the results reported in the literature are not homogenous.<sup>18, 23, 24</sup>

Due to differences in *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1* genotype frequencies among ethnic groups, evaluation of these genetic markers in the predisposition to drug-induced hepatotoxicity during TB treatment is needed, mainly in admixed populations. Brazilians are members of one of the most heterogeneous population in the world, result of five centuries of ethnic admixture of people from three continents: the European colonizers mainly represented by the Portuguese, the African slaves and the autochthonous Amerindians. Having in mind the diversity of those genes distribution and the high incidence of tuberculosis in Brazil, we have reported here the contribution of *NAT2*, *CYP2E1*, *GSTM1* and *GSTT1* genetic variants in susceptibility to anti-TB drug-induced hepatotoxicity in Brazilians.

## **Materials and Methods**

### **Study design and Patients**

In a case-control design, a total of 167 unrelated active tuberculosis patients who had been treated between 1998 and 2008 at the University Hospital Complex: Thoracic Institute/Clementino Fraga University Hospital from Federal University of Rio de Janeiro (UFRJ) (Brazil) was enrolled in this study. We retrospectively identified all the patients that developed antituberculosis drug-induced hepatotoxicity and/or acute hepatitis (cases) and the ones that showed no evidence of antituberculosis hepatic side effects (controls).



This study was approved by the ethics committee of the Oswaldo Cruz Foundation (FIOCRUZ) of Rio de Janeiro, where the samples were processed and analyzed, and written informed consent was obtained from all the enrolled patients.

All participants fulfilled the following inclusion criteria: (i) age above 18 years; (ii) diagnosis of active tuberculosis; (iii) treatment with anti-TB drugs regimens that include isoniazid at the usual dosage (400mg/day); (iv) normal baseline serum transaminases [alanine aminotransferase (ALT) and aspartate aminotransferase (AST)] before treatment. The exclusion criteria were: (i) positive serological testing for the human immunodeficiency virus (HIV), hepatitis B virus or hepatitis C virus; (ii) alcohol abuse intake; (iii) history of chronic liver disease; (iv) pregnancy and (v) refusal of signing informed written consent.

A volume of 3 mL of venous blood was collected from each volunteers and stored at –20°C. Genomic DNA was isolated from 100 µL of frozen whole blood using the FlexiGene DNA Kit (Qiagen Inc., USA), according to the manufacturer's specifications. After extraction, DNA samples were stored at –20°C.

#### **Liver function monitoring and diagnosis of drug-induced adverse effects**

Liver function tests, consisting of serum transaminases, total bilirubin and alkaline phosphatase measurements, were done once a month, during the patient's visit, by using commercial kits according to the manufacturer's instructions. The normal values of different tests were: ALT 30-65 IU/L; AST 15-37 IU/L; total bilirubin up to 1mg/dL and alkaline phosphatase 50-136 IU/L. Clinical and laboratory checkups were performed more frequently when hepatitis symptoms or abnormal serum transaminases levels were observed.

According to the American Thoracic Society criteria<sup>6</sup> and Brazilian Health Ministry, cases were divided in two subgroups: (1) drug induced hepatotoxicity: individuals who showed antituberculosis drug-induced hepatotoxicity, with an increase in serum transaminases (either AST or ALT) values higher than two times the upper limit of normal values ( $> 2xULN$ ) during treatment; and (2) drug induced hepatitis: individuals who showed antituberculosis drug-

induced hepatitis, with an increase in serum transaminases values higher than three times the upper limit of normal values ( $> 3 \times \text{ULN}$ ) when the recommendation is to interrupt the treatment.

### ***NAT2* genotyping and haplotype reconstruction**

Genotyping of the *NAT2* coding region and computational analysis were performed as described previously.<sup>38</sup> In brief, a 1093 bp fragment, which contains the entire coding region of *NAT2*, was amplified by PCR and subsequently sequenced on both DNA strands using the primers for amplification and two other internal primers<sup>38</sup> Sequencing was performed using ABI PRISM Big Dye Terminator v. 3.1 Kit (PE Applied BioSystems), according to the manufacturer's recommendations, on an ABI PRISM 3730 DNA Analyser (PE Applied BioSystems). The sequence analysis and SNPs identification were performed using SeqScape v.2.5 software (Applied BioSystems <http://www.appliedbiosystems.com>) with the sequence GenBank AY331807 (<http://www.ncbi.nlm.nih.gov/>) as the reference. To solve the haplotype pair of each patient and to define the acetylation status, haplotype reconstruction using the program PHASE v2.1.1 was carried out<sup>39, 40</sup> with the default model for recombination rate variation.<sup>41</sup> Five independent runs with 1000 interactions, 500 burn-in interactions and a thinning interval of 1 were performed. The best run that showed the maximum consistency across five runs was chosen. Additionally, the consistency of the results applying the PHASE algorithm repeatedly with default and varying values of number of interactions, burn-in interactions and the thinning interval was pursued.

### ***CYP2E1* genotypes identification**

To genotype the SNP -1053C>T in the *CYP2E1* gene promoter region and to determine the three genotypes (*\*1A/\*1A*; *\*1A/\*5* and *\*5/\*5*), a polymerase chain reaction (PCR) with the amplification primers CYP2E1F 5'CCA GTC GAG TCT ACA TTG TCA 3' and CYP2E1R 5' TTC ATT CTG TCT TCT AAC TGG 3' were performed followed by RFLP as previously described.<sup>42</sup> Briefly, after the amplification of a 413 bp DNA fragment flanking the -1053C>T SNP the PCR product was digested with *Rsa I* restriction enzyme. Evaluation of the fragment digestions was achieved by electrophoresis on 2% agarose gel followed by ethidium bromide

staining (0.5µg/mL).

### ***GSTT1* and *GSTM1* genotypes identification**

In order to determine the presence or absence of the *GSTM1* and *GSTT1* genes, two separate multiplex PCR reactions were used as reported previously, with minor modifications.<sup>43</sup>

<sup>44</sup> The *GSTM1* wild-type allele was identified by PCR method using primers M1 5' CTG CCC TAC TTG ATT GAT GGG 3' and M2 5' CTG GAT TGT AGC AGA TCA TGC 3' to obtain a 273-bp PCR product.<sup>43</sup> Co-amplification of a 536 bp fragment of β-globin gene using the primers: Beta1 5' GGT TGG CCA ATC TAC TCC CAG G 3' and Beta2 5' GCT CAC TCA GTG TGG CAA AG 3' was performed as an internal control. The PCR reaction mixture was carried out in a total volume of 50 µL containing 100ng of genomic DNA, 240 ng of each *GSTM1* primers and 10 pmoles for each β-globin primer, PCR Gold buffer (1x), 3 mM MgCl<sub>2</sub>, 0.1% Triton X-100, 0.2 mM of each dNTP and 2U Ampli Taq Gold DNA polymerase (PE Applied Biosystems). Amplification conditions consisted of an initial denaturing step of 95°C for 10 min. followed by 30 cycles of 95°C for 30 s, 64°C for 1 min., and 72°C for 6 min. The homozygous *GSTM1* null genotypes were detected by the absence of the 273 bp fragment.

For *GSTT1* genotyping and identification of the three genotypes (*GSTT1*\*A/*GSTT1*\*A, *GSTT1*\*A/*GSTT1*\*0 and *GSTT1*\*0/*GSTT1*\*0) a 1460 bp *GSTT1*\*0 specific fragment was co-amplified with a 460 bp fragment corresponding to *GSTT1*\*A allele in a single reaction tube using primers T1 5' CCA GCT CAC CGG ATC ATG GCC AG 3', T2 5' GAG ATG TGA GGA CCA GTA AGG AAG G 3', T3 5' CAG TTG TGA GCC ACC GTA CCC 3' and T4 5' GAG GGG GCC AGC AAC TAT CG 3'.<sup>44</sup> PCR was carried out in 50 µL reaction mix containing 150 ng genomic DNA, PCR buffer (1x), 100 ng of each primer, 1.5 mM MgCl<sub>2</sub>, 100 µg/mL BSA, 0.2 mM of dNTPs and 2U DNA polymerase (Biotools, Madrid/Spain). Amplification conditions consisted of a touchdown PCR with an initial denaturing step (94°C for 5 min.) followed by 20 cycles of 94°C for 1 min., 68°C for 1 min. (with a decrease of 0,5°C per cycle at the annealing temperature), and 72°C for 1 min. and 30 s. Additional 15 cycles of

94°C for 1 min., 58°C for 1 min. and 72°C for 1.5 min. were performed. Evaluation of the PCR amplicons was achieved by electrophoresis on 2% agarose gel followed by ethidium bromide staining.

### **Statistical analysis**

Data were expressed as the mean  $\pm$  standard deviation (SD) (for normal distribution) or as median with range (for non normal distribution). Expected genotype frequencies were calculated from respective single allele frequencies according to the Hardy-Weinberg equation. The observed and expected gene frequencies were compared using a  $\chi^2$  test to the Hardy-Weinberg equilibrium.

Qualitative variables were compared using the  $\chi^2$  test with Yates correction, or the Fisher exact test when appropriate. Quantitative variables were analyzed using the Student's *t* test, for normal distribution, or Mann-Whitney U-test (2 groups) and Kruskal-Wallis test (up to 2 groups) for non normal distribution. Odds ratios (OR) and confidence intervals (CI) were calculated to evaluate the susceptibility of *NAT2* haplotypes, genotypes and acetylation status to antituberculosis drug-induced hepatic adverse effects. Multivariate analysis was performed to evaluate the adjusted risk of antituberculosis drug-induced hepatic adverse effects in the different genotypes using a stepwise logistic regression analysis. All the analyses were carried out with the statistical software SPSS 12.0 for Windows (SPSS Inc., Chicago, IL). All statistical test were evaluated at 0.05 significance level.

### **Results**

Among 167 individuals enrolled in this study, 37 presented antituberculosis drug-induced hepatotoxicity where 26 were diagnosed with antituberculosis drug-induced hepatitis. No statistical differences in age, sex and basal liver function tests between cases and controls were observed (Tables 1 and 2). The median of ALT, AST and total bilirubin during treatment

were statistically different between cases (hepatotoxicity or hepatitis) and controls, showing a clinically significant liver injury among cases (Tables 1 and 2).

The *CYP2E1*, *GSTM1* and *GSTT1* genotypes under investigation were in Hardy-Weinberg equilibrium ( $p>0.05$ ). There were no significant differences in the frequencies of *CYP2E1*, *GSTM1* and *GSTT1* genotypes between cases and controls neither for hepatotoxicity nor for hepatitis analyses (Tables 1 and 2).

After *NAT2* sequencing and haplotype pair reconstruction, 14 *NAT2* alleles were identified, of which, 3 were new, firstly described here (Table 3). Although the distribution of *NAT2* alleles between cases and controls was not statistically different, the alleles with low activity were more frequent in cases, while the functional alleles were more frequent in the control group (Table 3). Genotypes of eight individuals (4.8%) could not be solved after the reconstruction by PHASE v2.1.1. Additionally, we identified other 10 individuals whom the acetylation status has not yet been characterized (Tables 4 and 5). Therefore, a total of 18 (10.8%) individuals were excluded for *NAT2* analysis.

Among 149 individuals whose acetylation status could be determined, 82 (55.03%) were slow acetylators, 60 (40.27%) were intermediate acetylators and only 7 (4.7%) individuals showed fast acetylation genotypes. The slow acetylators had higher median serum ALT and AST levels than fast/intermediate acetylators during the treatment (ALT:  $p=0.03$ ; AST:  $p=0.01$ ). None specific *NAT2* genotype was found to be associated to susceptibility of anti-TB drug-induced hepatotoxicity or hepatitis occurrence (Tables 4 and 5).

The risk of antituberculosis drug-induced hepatitis occurrence was significantly higher in slow acetylators than in fast/intermediate acetylators (22% versus 9%; OR, 2.86; CI 95%, 1.06-7.68;  $p = 0.04$ ). The same result was not observed when the group with antituberculosis drug-induced hepatotoxicity was compared to controls (Tables 4 and 5).

Variables such as age, sex, acetylation status, and *CYP2E1*, *GSTM1* and *GSTT1* genotypes were further analyzed in a multivariate logistic regression model. After adjustment, the slow acetylation was the only risk factor for both antituberculosis drug-induced

hepatotoxicity (OR, 2.62; 95% IC, 1.75-3.49;  $p = 0.03$ ) and hepatitis (OR, 3.59; 95% IC, 2.53-4.64;  $p = 0.02$ ). We have also performed similar analyses classifying individuals in three acetylation phenotypes (slow, intermediate and fast acetylators) and observed the same results for hepatotoxicity (OR, 2.63; IC 95%, 1.81-3.45;  $p=0.02$ ) or for hepatitis (OR, 3.47; IC 95%, 2.43-4.47;  $p=0.02$ ).

## Discussion

Among the first line therapeutic anti-TB drugs, isoniazid is the main drug associated with hepatotoxicity,<sup>45-47</sup> the other drugs related with this adverse reaction are reported in combination with isoniazid.<sup>48</sup> NAT2 is the main enzyme in the isoniazid metabolism and it is expressed in the liver. Like most studies conducted in different countries,<sup>16, 17, 19, 20</sup> our findings suggest that *NAT2* is one of the genetic factors responsible for predisposition to anti-TB drug-induced hepatotoxicity in Brazilian individuals. Additionally, the incidence of elevated levels of serum transaminases was significantly higher in slow acetylators than in fast/intermediate acetylators. These results are in agreement with the current hypothesis about the risk factor of hepatic side effects based on acetylation status. Indeed, slow acetylators acetylate not only isoniazid more slowly but also acetylhydrazine, the immediate precursor of toxic intermediates, to the harmless diacetylhydrazine.<sup>49, 50</sup> This protective acetylation is further suppressed by isoniazid competition. Therefore, slow acetylators may critically and indirectly increase the accumulation of toxic metabolites. In addition, another important route to generate toxic intermediates from isoniazid is by direct hydrolysis of isoniazid without acetylation.<sup>51</sup> This reaction produces hydrazine, which can induce hepatic injury.<sup>10, 51</sup> Pharmacokinetics studies showed that the serum concentration of hydrazine was significantly higher in slow acetylators than in fast acetylators, probably due to the high concentration of isoniazid. The high amount of isoniazid disposed by this pathway lead to enhanced hydrolysis to hydrazine, since the rate of metabolic conversion of isoniazid to acetylisoniazid is lower in slow than in fast acetylators.<sup>52, 53</sup>

All of these drug-disposal processes may support the finding that slow acetylators are prone to isoniazid-induced hepatitis.

Although few studies had been reported the absence of relationship between acetylation status and hepatotoxicity during TB treatment<sup>34-36</sup> or observed that fast acetylators could be susceptible to this kind of side effects,<sup>33</sup> some considerations must be done. In those studies, phenotyping methods to access the acetylation status of TB patients were used,<sup>34-36</sup> which may have less efficacy being influenced by several confounding factors, such as age, sex, diet, drug interactions, environmental factors, and co morbid illnesses. Here, genotyping of the *NAT2* gene were performed by sequencing, overcoming these problems in the classification of the acetylation status. Other aspect that should be taken into account is that in our study, patients were treated with 400 mg of isoniazid whereas 300 mg were prescribed in the study conducted by Vuilleumier et al.<sup>36</sup> Finally, here we were very rigorous with selection criteria of individuals to avoid confounding variables.

In our study, no association was found with specific *NAT2* haplotype and/or genotype related to susceptibility of anti-TB drug-induced hepatotoxicity occurrence. These findings may be due to a high diversity in the *NAT2* locus of Brazilian individuals, meaning that a group of defective *NAT2* alleles are involved with hepatic side effects instead of a specific one.

*CYP2E1*, one of the CYP enzymes, may convert acetylhydrazine into hepatotoxins affecting anti-TB drug-induced hepatotoxicity. Some studies have shown the relationship between *CYP2E1*\*5/\*5 genotype and the risk of anti-TB drug-induced hepatotoxicity.<sup>21, 36</sup> However, a recent study in a Korean population did not confirm those results.<sup>37</sup> In Brazilians, it has been shown that the homozygous mutant *CYP2E1*\*5/\*5 is rare (0.5%),<sup>54</sup> which differs from Taiwanese population (12%).<sup>21</sup> Indeed, we found only one individual with that genotype among 167 patients enrolled in this study (0.6%) with no association between *CYP2E1* genotypes and hepatic side effects occurrence. The discrepancy of our results may be due to differences in the frequencies of wild-type *CYP2E1*\*1A and mutant *CYP2E1*\*5 alleles among Brazilian individuals. Therefore, further studies, with larger sample size, should be carried out in

Brazilian population to better evaluate the relationship between *CYP2E1* genotypes and adverse reactions during TB treatment.

Finally, the frequencies of homozygous *GSTM1* and *GSTT1* null genotypes in this study were in accordance to other studies in a Brazilian population<sup>55</sup> showing the consistence of our data. Few studies observed the relationship between *GSTs* genotypes and antituberculosis drug-induced hepatotoxicity.<sup>18, 23, 24</sup> Indeed, Roy et al. have observed a significant association between the homozygous *GSTM1* null polymorphism and the antituberculosis drug-induced hepatotoxicity in Indian TB patients.<sup>18</sup> Recently, Huang et al. have found similar results in a Chinese population.<sup>23</sup> On the other hand, the presence of homozygous *GSTT1* null polymorphism was similar between cases and controls in both studies. The same analysis in Spanish TB patients however, showed an opposite effect: the homozygous *GSTT1* null polymorphism was highly associated with anti-TB drug-induced hepatotoxicity and no significant associations were found between homozygous *GSTM1* null and hepatotoxicity.<sup>24</sup> So, further studies should be carried out to evaluate these genetic factors in the susceptibility to hepatotoxicity occurrence.

Although in a previous study, the association of slow acetylation profile with higher risk of developing hepatotoxicity to anti-tuberculosis drugs has already been reported in Brazilians from the Southern region,<sup>57</sup> to our knowledge, this is the first study with Brazilian individuals where four genetic markers were analyzed to evaluate their relationship with side effects occurrence during antituberculosis therapy. The high diversity of *NAT2* gene and the presence of rare alleles in Brazilian individuals, identified after sequencing that could not be solved by PHASE program, lead to a reduced number of patients enrolled for *NAT2* analysis. Even so, critically considering and excluding the confounding factors, the results presented here are the first experimental proof to support the influence of the genetic factors in the occurrence of antituberculosis drug-induced liver adverse reactions among Brazilians and may be used to predict these side effects.



Therefore, the prediction of antituberculosis drug-induced hepatotoxicity would not only assist in determining the most suitable drug doses for individual patients, but should also reduce the number of outpatient visits, laboratory tests and hospitalizations due to liver dysfunction.

### **Acknowledgements**

We thank the nursing service of the Hospital Tuberculosis Control Program/Thoracic Diseases Institute/Clementino Fraga Filho University Hospital, for sample collection and the support from the PDTIS- FIOCRUZ sequencing and Bioinformatics Platforms.

This work was supported by Faperj/Pronex: Proc: E-26/170.0003/2008 and FAPERJ Pensa Rio: E-26/110.288/2007. Adalberto Rezende Santos, Afrânio Lineu Kritski, Fernanda C. Q. Mello, and Philip Noel Suffys are supported by CNPq Grant numbers 308786/2005-0, 35 04 77/1995-7, 306702/2007-0 and 312165/2006-4 respectively.

### **References**

1. World Health Organization. Global tuberculosis control. WHO Report 2008. World Health Organization, Geneva, Switzerland.
2. Santos-Filho ET. Public Health Watch. TB politics in Brazil – A perspective of Civil Society. Rio de Janeiro: Open Society Institute, 2006.
3. Health Ministry of Brazil. National politics of TB control. Brasília, Brazil; 2007.
4. Garibaldi RA, Drusin, RE, Ferebee, SH, Gregg MB: Isoniazid-associated hepatitis. Report of an outbreak. *Am Rev Respir Dis* 1972; **106**:357-365.
5. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C: Tuberculosis. *Lancet* 2003; **362**:887-899.
6. Saukkonen JJ, Cohn DL, Jasmer RM et al: An official ATS statement: hepatotoxicity of antituberculosis therapy. *Am J Respir Crit Care Med* 2006; **174**:935-952.

7. Mitchell JR, Snodgrass WR, Gillette JR: The Role of Biotransformation in Chemical-Induced Liver Injury. *Environ Health Perspect* 1976; **15**:27-38.
8. Nelson SD, Mitchell JR, Timbrell JA, Snodgrass WR, Corcoran GB: Isoniazid and iproniazid: activation of metabolites to toxic intermediates in man and rat. *Science* 1976; **193**:901-3.
9. Woodward KN, Timbrell JA: Acetylhydrazine hepatotoxicity: the role of covalent binding. *Toxicology* 1984; **30**:65-74.
10. Timbrell JA, Mitchell JR, Snodgrass WR, Nelson SD: Isoniazid hepatotoxicity: the relationship between covalent binding and metabolism in vivo. *J Pharmacol Exp Ther* 1980; **213**:364-369.
11. Lee, WM: Medical progress: drug-induced hepatotoxicity. *N Engl J Med* 2003; **349**:474-485.
12. Sodhi CP, Rana SV, Mehta SK et al: Study of oxidative stress in isoniazid-induced hepatic injury in young rats with and without protein-energy malnutrition. *J Biochem Toxicol* 1996, **11**:139-146.
13. Hayes JD, Flanagan JU, Jowsey IR: Glutathione transferases. *Annu Rev Pharmacol Toxicol* 2005; **45**:51-88.
14. Kimmoun E, Samuel D: Antituberculous drugs in patients with chronic liver disease. *J Gastroenterol hepatol* 2002; **17**:408-412.
15. Yew WW, Leung CC: Antituberculosis drugs and hepatotoxicity. *Respirology* 2006; **11**:699-707.
16. Ohno M, Yamaguchi I, Yamamoto I 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.
17. Kita T, Tanigawara Y, Chikazawa S et al: N-acetyltransferase 2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients. *Biol Pharm Bull* 2001; **24**:544-549.

18. Roy B, Chowdhury A, Kundu S et al: Increased risk of antituberculosis drug-induced hepatotoxicity in individuals with glutathione S-transferase M1 “null” mutation. *J Gastroenterol Hepatol* 2001; **16**:1033-1037.
19. Huang YS, Chern HD, Su WJ et al: Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor for antituberculosis drug-induced hepatitis. *Hepatology* 2002; **35**:883-889.
20. Hiratsuka M, Kishikawa Y, Takekuma Y et al: Genotyping the N-acetyltransferase 2 polymorphism in the prediction of adverse drug reactions to isoniazid in Japanese patients. *Drug Metabol Pharmacokin* 2002; **17**:357-362.
21. Huang YS, Chern HD, Su WJ et al: Cytochrome P450 2E1 genotype and the susceptibility to antituberculosis drug-induced hepatitis. *Hepatology* 2003; **37**:924-930.
22. Roy B, Ghosh SK, Sutradhar D, Sikdar N, Mazumder S, Barman S: Predisposition of antituberculosis drug induced hepatotoxicity by cytochrome P450 2E1 genotype and haplotype in pediatric patients. *J Gastroenterol Hepatol* 2006; **21**:784-786.
23. Huang YS, Su WJ, Huang YH et al: Genetic polymorphisms of manganese superoxide dismutase, NAD(P)H:quinine oxidoreductase, glutathione S-transferase M1 and T1, and the susceptibility to drug-induced liver injury. *J Hepatol* 2007; **47**:128-134.
24. LeiroV, Fernández-Villar A, Valverde D et al: Influence of glutathione S-transferase M1 and T1 homozygous null mutations on the risk of antituberculosis drug-induced hepatotoxicity in Caucasian population. *Liver Int* 2008; **28**:835-839.
25. Fretland AJ, Leff MA, Doll MA, Hein DW: Functional characterization of human N-acetyltransferase 2 (NAT2) single nucleotide polymorphisms. *Pharmacogenetics* 2001;**11**:207-215.
26. 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.

27. Parkin DP, Vandenplas S, Botha FJ et al: Trimodality of isoniazid elimination: phenotype and genotype in patients with tuberculosis. *Am J Respir Crit Care Med* 1997; **155**:1717-1722.
28. Kinzig-Schippers M, Tomalik-Scharte D, Jetter A et al: Should we use N-acetyltransferase type 2 genotyping to personalize isoniazid doses? *Antimicrob Agents Chemother* 2005; **49**:1733–1738.
29. Chen B, Zhang WX, Cai WM: The influence of various genotypes on the metabolic activity of NAT2 in Chinese population. *Eur J Clin Pharmacol* 2006; **62**:355-359.
30. Hickman D, Sim E: N-acetyltransferase polymorphism: Comparison of phenotype and genotype in humans. *Biochem Pharmacol* 1991; **42**:1007-1014.
31. Moussa LA, Khassouani CE, Hue B, Jana M, Begaud B, Soulaymani R: Determination of the acetylators phenotype in Moroccan tuberculosis patients using isoniazid as metabolic probe. *Int J Clin Pharmacol Ther* 2002; **40**:548-553.
32. Jetter A, Kinzig-Schippers M, Illauer M et al: Phenotyping of N-acetyltransferase type 2 by caffeine from uncontrolled dietary exposure. *Eur J Clin Pharmacol* 2004; **60**:17-21.
33. Mitchell JR, Thorgeisson UP, Black M et al: Increased incidence of isoniazid hepatitis in rapid acetylators: possible relation to hydralazine metabolites. *Clin Pharmacol Ther* 1975; **18**:70–79.
34. Singh J, Garg PK, Thakur VS, Tandon RK: Antitubercular treatment induced hepatotoxicity: does acetylator status matter? *Indian J Physiol Pharmacol* 1995; **39**:43–46.
35. Singh J, Arora A, Garg PK, Thakur VS, Pande JN, Tandon RK: Antituberculosis treatment-induced hepatotoxicity: role of predictive factors. *Postgrad Med J* 1995; **71**:359–362.
36. Vuilleumier N, Rossier MF, Chiappe A et al: CYP2E1 genotype and isoniazid-induced hepatotoxicity in patients treated for latent tuberculosis. *Eur J Clin Pharmacol* 2006; **62**:423-429.

37. Cho HJ, Koh WJ, Ryu YJ et al: Genetic polymorphisms of *NAT2* and *CYP2E1* associated with antituberculosis drug-induced hepatotoxicity in Korean patients with pulmonary tuberculosis. *Tuberculosis* 2007; **87**:551-556.
38. Teixeira RLF, Miranda AB., Pacheco AGF 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.
39. Stephens M, Smith NJ, Donnelly P: A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; **68**:978-989.
40. Stephens M, Donnelly P: A comparison of Bayesian methods for haplotype reconstruction. *Am J Hum Genet* 2003; **73**:1162-1169.
41. Li N, Stephens M: Modeling linkage disequilibrium, and identifying recombination hotspots using SNP data. *Genetics* 2003; **165**:2213-2233.
42. Stephens EA, Taylor JA, Kaplan N et al: Ethnic variation in the *CYP2E1* gene: polymorphism analysis of 695 African-Americans, European-Americans and Taiwanese. *Pharmacogenetics* 1994; **4**:185-192.
43. 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.
44. Sprenger R, Schlagenhauer R, Kerb R et al: Characterization of the glutathione S-transferase *GSTT1* deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype=phenotype correlation. *Pharmacogenetics* 2000; **10**:557-565.
45. Steele MA, Burk RF, Desprez RM: Toxic hepatitis with isoniazid and rifampin. A meta-analysis. *Chest* 1991; **99**:465-471.
46. Salpeter SR: Fatal isoniazid-induced hepatitis. Its risk during chemoprophylaxis. *West J Med* 1993; **159**:560-564.

47. Fountain FF, Tolley E, Chrisman CR, Self TH: Isoniazid hepatotoxicity associated with treatment of latent tuberculosis infection: a 7-year evaluation from a public health tuberculosis clinic. *Chest* 2005; **128**:116-123.
48. Tostmann A, Boeree MJ, Aarnoutse RE, de Lange WC, van der Ven AJ, Dekhuijzen R: Antituberculosis drug-induced hepatotoxicity: concise up-to-date review. *J Gastroenterol Hepatol* 2008; **23**:192-202.
49. Lauterburg BH, Smith CV, Todd EL, Mitchell JR: Pharmacokinetics of the toxic hydrazine metabolites formed from isoniazid in humans. *J Pharmacol Exp Ther* 1985; **235**:566-570.
50. Ellard GA, Gammon PT: Pharmacokinetics of isoniazid metabolism in man. *J Pharmacokinet Biop* 1976; **4**:83-113.
51. Timbrell JA, Wright JM, Baillie TA: Monoacetylhydrazine as a metabolite of isoniazid in man. *Clin Pharmacol Ther* 1977; **22**:602-608.
52. Sarma GR, Immanuel C, Kailasam S, Narayana AS, Venkatesan P: Rifampin-induced release of hydrazine from isoniazid. A possible cause of hepatitis during treatment of tuberculosis with regimens containing isoniazid and rifampin. *Am Rev Respir Dis* 1986; **133**:1072-1075.
53. Fukino K, Sasaki Y, Hirai S et al: Effects of NAT2, CYP2E1 and GST genotypes on the serum concentrations of isoniazid and metabolites in tuberculosis patients. *J Toxicol Sci* 2008; **33**:187-95.
54. Sigimura H, Hamada GS, Suzuki I et al: CYP1A1 and CYP2E1 polymorphism and lung cancer, case-control study in Rio de Janeiro, Brazil. *Pharmacogenetics* 1995; **5**:S145-148.
55. Suarez-Kurtz G, Vargens DD, Struchiner, CJ, Bastos-Rodrigues L, Pena SDJ: Self-reported skin color, genomic ancestry and the distribution of *GST* polymorphisms. *Pharmacogenet Genomics* 2007; **17**:765-771.

56. Roy PD, Majumder M, Roy B: Pharmacogenomics of anti-TB drugs-related hepatotoxicity. *Pharmacogenomics* 2008; **9**:311-321.
57. Possuelo LG, Castelan JA, de Brito TC et al: Association of slow N-acetyltransferase 2 profile and anti-TB drugs induced hepatotoxicity in patients from Southern Brazil. *Eur J Clin Pharmacol* 2008; **64**:673-681.

Table 1: Demographic, clinical and laboratory data of enrolled subjects. Influence of the genotype and allele frequencies of *CYP2E1*, *GSTT1* and *GSTM1* in drug-induced hepatotoxicity.

	Hepatotoxicity (>2xULN AST or ALT)		<i>p</i>
	Presence (n=37)	Absence (n=130)	
Age (year)*	46.7±18.91	42.85±14.85	0.258 <sup>a</sup>
Sex (M/F)	22/15	68/62	0.441 <sup>b</sup>
<b>Baseline</b>			
AST**	27 (14-73)	24 (10-58)	0.495 <sup>c</sup>
ALT**	30 (14-95)	28 (10-73)	0.145 <sup>c</sup>
Total bilirubin**	0.5 (0.1-1.0)	0.5 (0.2-1.0)	0.264 <sup>c</sup>
Alkaline phosphatase**	81 (43-174)	84.5 (10-210)	0.804 <sup>c</sup>
<b>During antituberculosis treatment</b>			
Peak AST**	263 (57-2,600)	27 (13-74)	< <b>0.001</b> <sup>c</sup>
Peak ALT**	180 (53-1,331)	32 (13-73)	< <b>0.001</b> <sup>c</sup>
Peak total bilirubin**	2.5 (0.2-22.6)	0.65 (0.4-1.8)	<b>0.003</b> <sup>c</sup>
Peak alkaline phosphatase**	152 (54-591)	175 (60-450)	0.922 <sup>c</sup>
<b><i>CYP2E1</i></b>			
*1A/*1 <sup>a</sup>	33 (89.2%)	118 (90.8%)	0.168 <sup>b</sup>
*1A/*5	3 (8.1%)	12 (9.2%)	
*5/*5	1 (2.7%)	0 (0%)	
*1 <sup>a</sup>	69 (93.2%)	248 (95.4%)	-
*5	5 (6.8%)	12 (4.6%)	-
<b><i>GSTT1</i></b>			
+/+	12 (32.4%)	34 (26.1%)	0.720 <sup>b</sup>
+/-	18 (48.7%)	72 (55.4%)	
-/-	7 (18.9%)	24 (18.5%)	
Functional allele (wild-type)	42 (56.8%)	140 (53.8%)	-
null allele	32 (43.2%)	120 (46.2%)	-
<b><i>GSTM1</i></b>			
+/_	18 (48.6%)	77 (59.2%)	0.251 <sup>b</sup>
-/-	19 (51.4%)	53 (40.8%)	

Normal Interval of AST: 15-37 U/L; ALT: 30-65 U/L; Total bilirubin: <1mg/dL; Alkaline phosphatase: 50-136 U/L;

\* Mean e standard deviation; \*\* Median (range).

<sup>a</sup> *p*-value of Student *t* test; <sup>b</sup> *p*-value of  $\chi^2$ ; <sup>c</sup> *p*-value of Mann Whitney test



Table 2: Demographic, clinical and laboratory data of enrolled subjects. Influence of the genotype and allele frequencies of *CYP2E1*, *GSTT1* and *GSTM1* in drug-induced hepatitis

	Hepatitis (>3xULN AST or ALT)		P
	Presence (n=26)	Absence (n=141)	
Age (year)*	47.58±20.07	42.99±14.94	0.275 <sup>a</sup>
Sex (M/F)	16/10	74/67	0.395 <sup>b</sup>
<b>Baseline</b>			
AST**	27 (14-54)	24 (10-73)	0.577 <sup>c</sup>
ALT**	31 (14-95)	28 (10-73)	0.441 <sup>c</sup>
Total bilirubin**	0.5 (0.2-1.0)	0.5 (0.1-1.0)	0.211 <sup>c</sup>
Alkaline phosphatase**	80.5 (43-174)	84.5 (10-210)	0.787 <sup>c</sup>
<b>During antituberculosis treatment</b>			
Peak AST**	351.5 (74-2,600)	28.5 (13-118)	< 0.001 <sup>c</sup>
Peak ALT**	254 (60-1,331)	34 (13-155)	< 0.001 <sup>c</sup>
Peak total bilirubin**	3.9 (0.6-22.6)	0.6 (0.2-7.5)	< 0.001 <sup>c</sup>
Peak alkaline phosphatase**	154 (54-591)	163 (60-450)	0.884 <sup>c</sup>
<b><i>CYP2E1</i></b>			
*1A/*1A	23 (88.5%)	128 (90.8%)	
*1A/*5	2 (7.7%)	13 (9.2%)	0.064 <sup>b</sup>
*5/*5	1 (3.8%)	0 (0%)	
*1A	48 (92.3%)	269 (95.4%)	-
*5	4 (7.7%)	13 (4.6%)	-
<b><i>GSTT1</i></b>			
+/+	9 (34.6%)	37 (26.2%)	
+/-	13 (50%)	77 (54.6%)	0.666 <sup>b</sup>
-/-	4 (15.4%)	27 (19.2%)	
Functional allele (wild-type)	31 (59.6%)	151 (53.5%)	-
null allele	21 (40.4%)	131 (46.5%)	-
<b><i>GSTM1</i></b>			
+/_	15 (41.7%)	80 (56.7%)	
-/-	11 (42.3%)	61 (43.3%)	0.928 <sup>b</sup>

Normal Interval of AST: 15-37 U/L; ALT: 30-65 U/L; Total bilirubin: <1mg/dL; Alkaline phosphatase: 50-136 U/L;

\* Mean ± standard deviation; \*\* Median (range).

<sup>a</sup> p-value of Student *t* test; <sup>b</sup> p-value of  $\chi^2$ ; <sup>c</sup> p-value of Mann Whitney test

Table 3: NAT2 alleles frequencies among Brazilians TB patients enrolled in this study.

Allele	Haplotype*	Phenotype	Total	HepTox <sup>1</sup>	HepMed <sup>2</sup>
NAT2*6A	GTTACAGAAGG	slow	242	284	288
NAT2*5B	GCCATGGAGGG	slow	23	28.4	26.9
NAT2*4	GCTACGGAAGG	fast	18.9	13.5	13.5
NAT2*5A	GCCATGGAAGG	slow	12	9.5	9.6
NAT2*7B	GTTACGGAAGA	slow	4.5	6.8	7.7
NAT2*14B	ATTACGGAAGG	slow	3	2.7	3.8
NAT2*12A	GCTACGGAGGG	fast	2.4	0	0
NAT2*5C	GCCACGGAGGG	slow	2.7	2.7	3.8
NAT2*13A	GTTACGGAAGG	fast	1.5	1.3	0
NAT2*6H	GTTACAGGAGG	ND	0.9	2.7	1.9
NAT2*12G	GCTACGTAGGG	ND	0.6	1.3	0
NAT2* new1	GTTACAGAAAG	ND	0.9	0	0
NAT2* new2	GTTATGGAAGA	ND	0.3	0	0
NAT2* new3	GCCCTGGAGGG	ND	0.3	0	0
ND	-	ND	4.8	2.7	3.8
<i>Alleles with low activity</i>		-	69.5	78.4	80.8
<i>Functional alleles</i>		-	22.7	14.9	13.4
<i>Function not determined</i>		-	7.8	6.7	5.8
<b>Total</b>			<b>100</b>	<b>100</b>	<b>100</b>

\* Nucleotide positions within NAT2 coding region: 191; 282; 341; 472; 481; 590; 609; 766; 803; 838; 857;

<sup>1</sup> Allele frequency among individuals with hepatotoxicity (n° of individuals=37);

<sup>2</sup> Allele frequency among individuals with hepatitis (n° of individuals=26);

<sup>ND</sup> not determined.

Table 4: *NAT2* genotype/phenotype and susceptibility of isoniazid-induced hepatotoxicity.

Genotype/Phenotype	Hepatotoxicity		Total	OR*	95% CI*
	Presence	Absence			
<b>Slow acetylator</b>	<b>23</b> (62.2%)	<b>59</b> (45.4%)	<b>82</b> (49.1%)	2.22	0.97-5.08
5B/5B	7 (18.9%)	15 (11.5%)	22 (13.1%)	1.79	0.67-4.78
6A/6A	5 (13.5%)	10 (7.7%)	15 (9.9%)	1.88	0.60-5.88
6A/5A	2 (5.4%)	11 (8.5%)	13 (7.8%)	0.62	0.13-2.92
6A/5B	3 (8.1%)	5 (3.8%)	8 (4.8%)	2.21	0.50-9.68
7B/5A	1 (2.7%)	3 (2.3%)	4 (2.4%)	-	-
5C/5B	0 (0%)	4 (3.1%)	4 (2.4%)	-	-
6A/7B	0 (0%)	3 (2.3%)	3 (1.8%)	-	-
7B/5B	1 (2.7%)	2 (1.5%)	3 (1.8%)	-	-
5C/5C	1 (2.7%)	1 (0.8%)	2 (1.2%)	-	-
5A/5B	1 (2.7%)	1 (0.8%)	2 (1.2%)	-	-
5A/14B	0 (0%)	2 (1.5%)	2 (1.2%)	-	-
6A/14B	2 (5.4%)	0 (0%)	2 (1.2%)	-	-
7B/14B	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
5B/14B	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
<b>Intermediate acetylator</b>	<b>10</b> (27%)	<b>50</b> (38.4%)	<b>60</b> (35.9%)	-	-
6A/4	4 (10.8%)	16 (12.3%)	20 (12%)	0.86	0.27-2.76
4/5A	2 (5.4%)	13 (10%)	15 (9.9%)	0.51	0.11-2.39
4/5B	1 (2.7%)	7 (5.4%)	8 (4.8%)	-	-
12A/5B	0 (0%)	4 (3.1%)	4 (2.4%)	-	-
4/14B	0 (0%)	3 (2.3%)	3 (1.8%)	-	-
7B/4	2 (5.4%)	1 (0.8%)	3 (1.8%)	-	-
6A/13A	0 (0%)	2 (1.5%)	2 (1.2%)	-	-
6A/12A	0 (0%)	2 (1.5%)	2 (1.2%)	-	-
13A/5B	1 (2.7%)	1 (0.8%)	2 (1.2%)	-	-
13A/5A	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
<b>Fast acetylator</b>	<b>0</b> (0%)	<b>7</b> (5.4%)	<b>7</b> (4.2%)	-	-
4/4	0 (0%)	6 (4.6%)	6 (4.6%)	-	-
12A/12A	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
<b>Acetylation not determined</b>	<b>4</b> (10.8%)	<b>14</b> (10.8%)	<b>18</b> (10.8%)	-	-
6H/4	1 (2.7%)	0 (0%)	1 (0.6%)	-	-
12G/5C	0 (0%)	1 (0.8%)	1 (0.6%)	-	-

6H/5A	1 (2.7%)	0 (0%)	1 (0.6%)	-	-
6A/6H	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
7B/12G	1 (2.7%)	0 (0%)	1 (0.6%)	-	-
NEW1/14B	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
NEW2/5A	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
NEW1/4	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
NEW1/5A	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
5B/NEW3	0 (0%)	1 (0.8%)	1 (0.6%)	-	-
Genotype not determined	1 (2.7%)	7 (5.4%)	8 (4.8%)	-	-
<b>Total</b>	<b>37 (100%)</b>	<b>130 (100%)</b>	<b>167 (100%)</b>	-	-

\* Intermediate and fast acetylators were grouped into the same class and individuals with acetylation status not determined were excluded of phenotype analysis. Other genotypes were used as reference for OR.

Table 5: *NAT2* genotype/phenotype and susceptibility of isoniazid-induced hepatitis.

Genotype/Phenotype	Hepatitis		Total	OR*	95% CI*
	Presence	Absence			
<b>Slow acetylator</b>	<b>18</b> (69.2%)	<b>64</b> (45.4%)	<b>82</b> (49%)	<b>2.86</b>	<b>1.06-7.68</b>
5B/5B	5 (19.2%)	17 (12.1%)	22 (13.1%)	1.74	0.58-5.21
6A/6A	3 (11.5%)	12 (8.5%)	15 (8.9%)	1.40	0.37-5.36
6A/5A	2 (7.7%)	11 (7.8%)	13 (7.8%)	0.98	0.21-4.73
6A/5B	3 (11.5%)	5 (3.6%)	8 (4.8%)	3.55	0.79-15.87
7B/5A	1 (3.8%)	3 (2.1%)	4 (2.4%)	-	-
5C/5B	0 (0%)	4 (2.8%)	4 (2.4%)	-	-
6A/7B	0 (0%)	3 (2.1%)	3 (1.8%)	-	-
7B/5B	1 (3.8%)	2 (1.4%)	3 (1.8%)	-	-
5C/5C	1 (3.8%)	1 (0.7%)	2 (1.2%)	-	-
5A/5B	0 (0%)	2 (1.4%)	2 (1.2%)	-	-
5A/14B	0 (0%)	2 (1.4%)	2 (1.2%)	-	-
6A/14B	2 (7.7%)	0 (0%)	2 (1.2%)	-	-
7B/14B	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
5B/14B	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
<b>Intermediate acetylator</b>	<b>6</b> (23.1%)	<b>54</b> (38.3%)	<b>60</b> (36%)	-	-
6A/4	2 (7.7%)	18 (12.8%)	20 (12%)	0.57	0.12-2.62
4/5A	2 (7.7%)	13 (9.2%)	15 (9%)	0.82	0.17-3.87
4/5B	0 (0%)	8 (5.7%)	8 (4.8%)	-	-
12A/5B	0 (0%)	4 (2.8%)	4 (2.4%)	-	-
4/14B	0 (0%)	3 (2.1%)	3 (1.8%)	-	-
7B/4	2 (7.7%)	1 (0.7%)	3 (1.8%)	-	-
6A/13A	0 (0%)	2 (1.4%)	2 (1.2%)	-	-
6A/12A	0 (0%)	2 (1.4%)	2 (1.2%)	-	-
13A/5B	0 (0%)	2 (1.4%)	2 (1.2%)	-	-
13A/5A	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
<b>Fast acetylator</b>	<b>0</b> (0%)	<b>7</b> (5%)	<b>7</b> (4.2%)	-	-
4/4	0 (0%)	6 (4.3%)	6 (3.6%)	-	-
12A/12A	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
<b>Acetylation not determined</b>	<b>2</b> (7.7%)	<b>16</b> (11.3%)	<b>18</b> (10.8%)	-	-
6H/4	1 (3.85%)	0 (0%)	1 (0.6%)	-	-
12G/5C	0 (0%)	1 (0.7%)	1 (0.6%)	-	-

6H/5A	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
6A/6H	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
7B/12G	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
NEW1/14B	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
NEW2/5A	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
NEW1/4	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
NEW1/5A	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
5B/NEW3	0 (0%)	1 (0.7%)	1 (0.6%)	-	-
Genotype not determined	1 (3.85%)	7 (5%)	8 (4.8%)	-	-
<b>Total</b>	<b>26 (100%)</b>	<b>141 (100%)</b>	<b>167 (100%)</b>	<b>-</b>	<b>-</b>

\* Intermediate and fast acetylators were grouped into the same class and individuals with acetylation status not determined were excluded of phenotype analysis. Other genotypes were used as reference for OR.



## ***4- DISCUSSÃO***



Variações genéticas em enzimas metabolizadoras de drogas são responsáveis por diferenças interindividuais na resposta terapêutica, na suscetibilidade a doenças, na ocorrência de efeitos colaterais e interações medicamentosas. Assim, a identificação de marcadores genéticos que possam prever desfechos terapêuticos desfavoráveis com uma determinada droga pode direcionar não somente a indicação do medicamento correto como também do esquema posológico mais adequado, maximizando a eficácia e prevenindo ou reduzindo as possíveis reações adversas tóxicas.

Durante muitos anos a isoniazida vem sendo apontada como uma droga hepatotóxica e alguns estudos têm demonstrado uma associação entre o fenótipo/genótipo de acetilação de NAT2 e a suscetibilidade à ocorrência dessas reações adversas. Alguns polimorfismos no gene de NAT2 estão relacionados a uma baixa atividade enzimática, o que leva a geração de metabólitos hepatotóxicos a partir da biotransformação da isoniazida. Em contrapartida, a presença de dois alelos funcionais caracteriza um fenótipo de acetilação rápida, podendo interferir na eficácia da droga, o que torna a análise do perfil de acetilação um ponto crucial para a prevenção de desfechos terapêuticos desfavoráveis no tratamento anti-TB. Hoje, a genotipagem de NAT2 tem sido considerada um método rápido e simples para inferir o fenótipo de acetilação num curto espaço de tempo. Através desse acompanhamento genético, terapias individualizadas para drogas que são metabolizadas por NAT2 poderiam ser introduzidas na prática clínica, através do ajuste da dosagem do medicamento ou da administração de uma droga alternativa, de forma a maximizar o efeito terapêutico e reduzir as reações adversas no indivíduo sob tratamento.

Polimorfismos no gene que codifica para N-acetiltransferase 2 tem sido extensivamente investigados; hoje, a variabilidade do gene NAT2 em diferentes populações é um consenso geral. Uma vez que, ao iniciarmos esse estudo, não encontramos dados na literatura sobre o perfil de acetilação da população brasileira, bem como quais alelos circulavam aqui no país, iniciamos um mapeamento da região codificante de NAT2 de indivíduos residentes nos Estados do Rio de Janeiro e Goiás, utilizando como estratégia de genotipagem o sequenciamento direto de produto de PCR (Aprovação Comitê de Ética em Pesquisa n° 216/03 – Anexo 1).

Para tal, desenhamos quatro iniciadores, baseados na sequência referência de NAT2 (AY331807) contida no GenBank. Esses iniciadores foram utilizados tanto para a amplificação da região de interesse, como para as reações de sequenciamento (Anexo 2). Uma DNA polimerase com atividade revisora (*Tli* DNA polimerase) foi adicionada à reação de PCR para evitar a introdução de nucleotídeos de forma errônea pela *Taq* DNA polimerase, e





assim nos certificarmos da fidelidade da sequência amplificada. Dessa forma, conseguimos obter quatro sequências para cada amostra, a partir do produto amplificado de 1093 pb que, depois de editadas através de programa computacional apropriado, foram capazes de nos fornecer a sequência nucleotídica completa da região que codifica para NAT2. De modo a aperfeiçoar as análises das sequências obtidas a partir de quatrocentos e quatro indivíduos, o *software* SeqScape<sup>®</sup> foi utilizado nesse estudo e se mostrou apropriado como ferramenta para a identificação de SNPs já descritos, bem como para a validação de novos SNPs (Anexo 2).

Através da genotipagem por sequenciamento, conseguimos não somente determinar as frequências dos SNPs já descritos, como também identificar novas mutações em NAT2 e sugerir qual seria o perfil de acetilação das populações estudadas através da identificação dos genótipos encontrados na nossa amostragem. Um total de vinte e três SNPs foram identificados, sendo treze SNPs já descritos em outras populações e dez mutações, até o momento, ainda não descritas na literatura.

Com o objetivo de utilizarmos as novas variantes gênicas de NAT2 como ferramenta tecnológica para a predição de efeitos colaterais, falência terapêutica ou suscetibilidade a doenças onde este gene possa estar envolvido, redigimos um pedido de patente (Capítulo 2 - Resultados) solicitando a proteção de um método diagnóstico contendo as novas mutações encontradas em indivíduos brasileiros. Entretanto, durante a fase de redação, Patin e colaboradores (2006) publicaram dois artigos revelando a presença de quatro (472A>C, 578C>T, 622T>C, e 766A>G) das dez novas mutações na população africana (Patin *et al.*, 2006a; Patin *et al.*, 2006b). Sendo assim, a referida patente, depositada no âmbito nacional no dia 12 de outubro de 2007, e posteriormente internacionalizada em outubro de 2008 (Anexo 3), inclui apenas seis mutações (29T>C, 152G>T, 203G>A, 228C>T, 458C>T e 600A>G) descritas somente na fração da população brasileira estudada (Capítulo 2; Tabela 1 do Capítulo 3).

Dentre os SNPs já descritos, os sete mais comumente encontrados em diferentes regiões do mundo puderam ser observados numa frequência relevante em nossa população (Tabela 1 do Capítulo 1). Dentre esses polimorfismos, foi possível identificar mutações consideradas étnicoespecíficas, como os SNPs 191G>A, 341T>C e 857G>A. Com relação a influência dos SNPs no fenótipo de acetilação, observamos que quatro desses polimorfismos (191G>A, 341T>C, 590G>A e 857G>A) eram associados com um fenótipo de acetilação lenta e os outros três (282C>T, 481C>T e 803A>G) aparentemente não interferiam na atividade enzimática (Fretland *et al.*, 2001; Zang *et al.*, 2007). Seis mutações (70T>A; 345C>T; 403C>G; 609G>T; 683C>T e 838G>A), que não fazem parte do site oficial de



nomenclatura de *NAT2*, foram encontradas em frequências abaixo de 1%, e sua influência no fenótipo de acetilação ainda não foi determinada.

Em relação à origem étnicoespecífica dos SNPs, a mutação 191G>A foi inicialmente identificada em americanos-africanos através de um estudo realizado por Bell e colaboradores (1993). Esses autores sugeriram que o referido SNP teria uma origem africana, uma vez que era completamente ausente em populações eurodescendentes. (Bell *et al.*, 1993). Posteriormente, em 1996, outro estudo com populações africanas confirmou a origem do mesmo SNP (Delomenie *et al.*, 1996). Adicionalmente, diversos estudos mostram que a frequência de 191G>A pode variar nas diferentes regiões da África, com um decréscimo de sua incidência a partir do oeste para o leste africano, ao longo da África Subsaariana (Cavaco *et al.*, 2003). Nossos resultados mostram que a população de estudo apresenta uma frequência significativa (4,1%) de 191G>A, ficando atrás apenas das populações onde essa mutação se originou (Tabela 4 do Capítulo 1). O fato de populações do oeste da África terem participado do povoamento do Brasil na época da colonização talvez seja a melhor explicação para a presença deste SNP em nossa população (Parra *et al.*, 2003).

Por outro lado, o SNP 857G>A foi originalmente identificado em indivíduos japoneses (Ohsako e Deguchi, 1990), sendo raramente observado em populações européias (Blum *et al.*, 1991). Em 1993, Lin e colaboradores observaram que o SNP 857G>A caracterizava de 30-33% dos alelos lentos dentre os coreanos e chineses; entretanto, o mesmo caracteriza apenas 1-5% dos alelos lentos em populações eurodescendentes, corroborando a hipótese de que esta mutação originou-se na Ásia (Lin *et al.*, 1993). Estudos posteriores em populações ameríndias mostraram que esse SNP compreende a mutação mais frequente, dentre aqueles responsáveis pela diminuição na atividade de *NAT2* (Jorge-Nebert *et al.*, 2002) contribuindo assim para a hipótese da proximidade filogenética entre as populações asiática e ameríndia. Em nosso trabalho, o SNP 857G>A foi observado numa frequência de 4,4%, mostrando mais uma vez a participação de diferentes etnias na constituição da população brasileira (Tabela 4 do Capítulo 1).

Finalmente, quando analisamos a mutação 341T>C (considerada como de origem européia), pudemos verificar que a mesma se apresenta numa frequência similar às encontradas em populações eurodescendentes (Tabela 4 do Capítulo 1), mostrando, assim, a influência européia no perfil genético de *NAT2* da fração estudada da população brasileira.

No que diz respeito à diversidade alélica de *NAT2*, neste estudo inicial, fomos capazes de identificar dezesseis alelos diferentes de *NAT2* dos quais cinco deles ainda não foram caracterizados quanto ao fenótipo de acetilação (Tabela 2 do Capítulo 1). Dentre os onze



alelos com função conhecida, oito deles caracterizam um fenótipo de acetilação lenta, sendo os alelos *NAT2\*5B* (33%) e *NAT2\*6A* (26%) os mais frequentes. O alelo tipo-selvagem *NAT2\*4* foi o alelo funcional mais frequente em nossa amostragem (20%) e, quando comparamos a distribuição dos diferentes alelos nas duas regiões estudadas apenas este apresentou uma frequência significativamente diferente, sendo mais observado no Estado de Goiás (Tabela 2 do Capítulo 1). Essa distribuição alélica de *NAT2* mostra mais uma vez a forte influência de eurodescendentes e africanos na formação do *background* genético do povo brasileiro. De fato, em eurodescendentes e africanos a incidência de alelos do *cluster NAT2\*5* é alta, enquanto que a frequência dos alelos *NAT2\*7A* e *NAT2\*7B* é baixa (Mrozikiewicz *et al.*, 1996; Dandara *et al.*, 2003); já em populações asiáticas acontece o contrário, sendo os alelos dos *clusters \*6* e *7\** os responsáveis pelo fenótipo de acetilação lenta (Sekine *et al.*, 2001; Kukongviriyapan *et al.*, 2003). Em populações ameríndias, o alelo *NAT2\*7B* é o principal responsável pelo fenótipo de acetilação lenta, apresentando pouca ou nenhuma incidência de outros alelos que conferem uma baixa atividade de N-acetilação (Jorge-Nebert *et al.*, 2002).

Após a análise da combinação dos diferentes SNPs encontrados através do sequenciamento direto, conseguimos identificar o haplótipo de cento e vinte indivíduos homocigotos para todas as mutações ou heterocigotos para somente um SNP. Entretanto, após a exclusão de seis SNPs que não influenciavam no fenótipo de acetilação, pudemos caracterizar o genótipo de um total de duzentos e dois indivíduos e assim prever seu perfil de acetilação. Sendo assim, foi possível identificar quatorze genótipos e sugerir que o perfil de acetilação da população estudada é lento (49%), representado principalmente pelos genótipos *NAT2\*5/\*5* e *NAT2\*6/\*6* (Tabela 3 do Capítulo 1). Por outro lado, com base nas análises realizadas, o genótipo de acetilação rápida, nessas duas regiões do Brasil, é constituído pelos alelos *NAT2\*4*, *\*12* e *\*13*, semelhante ao observado em populações africanas, enquanto que em eurodescendentes a acetilação rápida é dependente principalmente da presença do alelo tipo-selvagem *NAT2\*4* (Cascorbi *et al.*, 1995; Delomenie *et al.*, 1996; Agundez, 2003).

Quando observamos separadamente os resultados obtidos nos dois Estados do Brasil, pudemos verificar que a população de Goiás apresenta uma maior prevalência de alelos funcionais quando comparado à população do Rio de Janeiro, além de conter uma diversidade alélica menor (Tabela 2 do Capítulo 1). Esse quadro pode ser explicado pela influência de populações ameríndias naquela região, uma vez que são populações mais homogêneas e com alto índice de alelos funcionais, representados pelo alelo *NAT2\*4* (Jorge-Nebert *et al.*, 2002).



Além disso, as duas populações estudadas apresentaram diferenças na distribuição dos genótipos de *NAT2* e perfil de acetilação, com uma frequência significativamente maior de acetiladores intermediários no Rio de Janeiro (Tabela 3 do Capítulo 1), mostrando a grande diversidade existente entre essas duas regiões do país.

A presença isolada de SNPs responsáveis por uma diminuição da atividade enzimática no gene que codifica para *NAT2* é um evento raro. Assim, quando duas ou mais destas mutações estão presentes numa determinada amostra, é de extrema importância averiguar se as mesmas estão localizadas num único alelo (sendo, o outro um alelo funcional), ou ainda se os SNPs estão presentes em alelos diferentes (caracterizando dois alelos não funcionais). Para se aferir o perfil de acetilação de um determinado indivíduo, é necessária a identificação do haplótipo de cada alelo de *NAT2*, uma vez que a atividade enzimática é influenciada pela combinação dos SNPs presentes em ambos os alelos.

A fim de caracterizar os alelos/haplótipos de *NAT2* portadores de mutações novas e/ou raras circulantes em indivíduos residentes no Rio de Janeiro e Goiás, partimos para a utilização da tecnologia de clonagem e ressequenciamento gênico (Anexo 4). Através dessa metodologia, conseguimos isolar os alelos de cada indivíduo analisado e identificar as combinações de mutações que geram os haplótipos, pelo sequenciamento dos DNAs recombinantes. Sete novos alelos foram identificados na fração da população brasileira estudada, apresentando diferentes frequências (Figura 1 do Capítulo 3).

Vários estudos de predição por modelagem molecular têm caracterizado a funcionalidade dos alelos de *NAT2* e obtido resultados semelhantes aos estudos experimentais (Rodrigues-Lima & Dupret, 2002; Zang *et al.*, 2007; Walraven *et al.*, 2008). Sendo assim, através de modelagem e análise estrutural da proteína *NAT2*, fomos capazes de sugerir o papel estrutural e/ou funcional de diferentes polimorfismos: encontramos uma mutação possivelmente envolvida no reconhecimento do substrato pela enzima (152G>T ou Gly51Val), uma mutação que altera o sítio catalítico pela perda de um grupamento funcional (203G>A ou Cys68Tyr) e mutações que podem levar à presença de resíduos hidrofóbicos expostos na superfície da proteína (458C>T, 578C>T, 683C>T, e 838G>A), facilitando a ubiquitilação e a degradação proteossômica da proteína (Grant, 2008). Todos esses polimorfismos parecem interferir na atividade de acetilação contribuindo ainda mais para o fenótipo de acetilação lenta predominante na fração estudada da população brasileira. Adicionalmente, dentre os sete novos alelos caracterizados pelo exame do DNA, cinco deles apresentaram potencial para codificar uma enzima com atividade reduzida (*NAT2\*5N*, *\*5O*, *\*6H*, *\*6M*, *\*12E*), outros dois (*NAT2\*20* e *\*12I*) parecem ser alelos funcionais por conterem



mutações silenciosas e/ou mutação que não altera a atividade enzimática (803A>G). Finalmente, os alelos *NAT2\*12F* e *NAT2\*12J* possivelmente codificam enzimas com atividade normal, uma vez que apresentam polimorfismos (622T>C ou 29T>C, em adição a 803A>G, respectivamente) que não afetam a estrutura ou a função enzimática.

Até o presente momento, este parece ser o primeiro trabalho de mapeamento da região codificante de *NAT2* por sequenciamento numa população miscigenada onde a tuberculose se apresenta de forma endêmica, servindo de base para estudos clínicos e epidemiológicos de *NAT2* no Brasil. Os diferentes SNPs encontrados, novos e já descritos, assim como a diversidade alélica para o gene *NAT2* mostram a heterogeneidade brasileira e isto se deve ao fato de diferentes grupos étnicos terem participado na construção de nossa população na época da colonização. Entretanto, a frequência esperada desses SNPs depende da região a ser estudada, devido a uma maior ou menor influência de determinado grupo étnico. Portanto, se faz necessário um estudo da prevalência dessas mutações e de padrões genotípicos em diferentes regiões do Brasil, para podermos avaliar com precisão as condutas terapêuticas do tratamento da tuberculose aplicadas em nosso país. Além disso, um estudo funcional da influência desses novos SNPs na atividade de *NAT2* seria importante para um completo entendimento e caracterização do fenótipo de acetilação no Brasil.

A fim de verificar a influência de variantes gênicas dos *loci NAT2, CYP2E1, GSTM1* e *GSTT1* na ocorrência de efeitos colaterais hepáticos induzidos por drogas anti-TB, realizamos um estudo retrospectivo do tipo caso-controle com pacientes atendidos no Hospital Universitário Clementino Fraga Filho, da Universidade Federal do Rio de Janeiro (Aprovação Comitê de Ética em Pesquisa n° 421/07 – Anexo 5). Para isso, os quatro genes foram genotipados por sequenciamento direto ou PCR, aliada, ou não, à técnica de RFLP (Anexos 2 e 6), e as frequências das variantes gênicas foram comparadas entre pacientes que desenvolveram ou não hepatotoxicidade durante o tratamento anti-TB com um esquema contendo INH.

Nossos resultados sugerem que os polimorfismos no gene *NAT2* constituem um fator de risco genético importante para a ocorrência de hepatotoxicidade induzida por drogas anti-TB. Através da regressão logística multivariada, fomos capazes de constatar que os acetiladores lentos apresentaram um risco significativamente maior em desenvolver hepatotoxicidade (OR, 2.62; IC 95%, 1.75-3.49;  $p = 0.03$ ) ou hepatite medicamentosa (OR, 3.59; IC 95%, 2.53-4.64;  $p = 0.02$ ) quando comparados aos acetiladores intermediários/rápidos (Capítulo 4). Além disso, observamos que aqueles indivíduos portadores de um genótipo/fenótipo de acetilação lenta apresentaram picos de transaminases



hepáticas significativamente mais altos quando comparados aos acetiladores intermediários/rápidos reforçando a hipótese de que indivíduos com uma deficiência na atividade de acetilação estariam mais predispostos a desenvolverem efeitos colaterais hepáticos durante o tratamento anti-TB.

De fato, vários estudos realizados em diferentes populações já constataram a existência de associação entre o fenótipo de acetilação lenta com a hepatotoxicidade induzida pela INH (Ohno *et al.*, 2000; Huang *et al.*, 2002; Hiratsuka *et al.*, 2002; Kinzig-Schippers *et al.*, 2005; Shimizu *et al.*, 2006; Cho *et al.*, 2007; Higuchi *et al.*, 2007). Através da genotipagem pela técnica de PCR-RFLP, esses autores comprovaram que, dentre os pacientes submetidos ao tratamento quimioterápico anti-TB contendo isoniazida, aqueles que continham dois alelos para acetilação lenta apresentavam um maior risco de desenvolvimento de hepatotoxicidade.

Alguns estudos não observaram uma relação entre o perfil de acetilação e a ocorrência de hepatotoxicidade (Singh *et al.*, 1995a; Singh *et al.*, 1995b; Roy *et al.*, 2006; Vuilleumier *et al.*, 2006), ou verificaram que os indivíduos com um fenótipo de acetilação rápida é que seriam mais suscetíveis à lesão hepática induzida por drogas anti-TB (Mitchell *et al.*, 1975). Várias possibilidades poderiam explicar as diferenças encontradas nesses trabalhos. Inicialmente, em alguns desses estudos a fenotipagem foi utilizada como ferramenta de associação entre o perfil de acetilação e o desenvolvimento de doença hepática. Entretanto, o fenótipo é influenciado por vários fatores, incluindo idade, sexo, hábitos alimentares, doenças, quimioterapia, consumo de álcool, contaminantes ambientais, etc. Além disso, dependendo do substrato utilizado como droga-sonda e do método analítico empregado, os fenótipos de acetilação muitas vezes não são distinguíveis, exibindo sobreposições. Sendo assim, os métodos de genotipagem passaram a ser mais utilizados, e a técnica mais comum para a identificação dos alelos de *NAT2* é o PCR-RFLP. Além de tratar-se de uma técnica simples, a utilização de genotipagem por PCR-RFLP para a determinação do perfil de acetilação apresenta outras vantagens quando comparada aos métodos de fenotipagem: é menos invasiva (não requer a administração de drogas evitando toxicidade para o indivíduo), é mais barata, mais rápida e reprodutível em larga escala, e não requer a supervisão do indivíduo, bastando apenas a coleta de alíquotas de sangue de onde o DNA genômico será extraído.

Outro elemento que pode causar erros de interpretação nesse tipo de estudo é a distribuição dos diferentes perfis de acetilação na população a ser estudada. Segundo Ohno e colaboradores (2000), a existência de poucos acetiladores rápidos com uma predominância de acetiladores intermediários impossibilita a identificação de diferenças na hepatotoxicidade



entre indivíduos lentos e rápidos. Outras duas hipóteses levantadas por esses autores seria que: (i) a frequência da administração do fármaco (diariamente ou com intervalos) poderia influenciar o acúmulo de substâncias tóxicas a partir da metabolização da INH e, (ii) a não exclusão de pacientes com fatores que os predispõe a hepatotoxicidade, como doenças no fígado, o consumo de álcool e a coadministração de outras drogas hepatotóxicas (Ohno *et al.*, 2000).

Outro fato importante que deve ser considerado é a diferença na posologia para a administração de INH utilizada na terapia da tuberculose nas diferentes populações de estudo. Por exemplo, no Japão a indicação é de 8mg/Kg/dia (até 400mg/dia) de INH, semelhante à utilizada na população aqui estudada, por outro lado, em populações eurodescendentes, a quimioterapia com essa droga consiste na administração de 5mg/Kg/dia (até 300mg/dia). Nesse caso, se os acetiladores lentos não apresentam um risco maior de desenvolvimento de lesão hepática na população eurodescendente, os resultados obtidos por Ohno e colaboradores indicam que a dosagem padrão administrada na população japonesa é alta para pessoas com um fenótipo de acetilação lenta (Ohno *et al.*, 2000).

Os métodos de genotipagem oferecem hoje uma alternativa simples e segura para a predição do fenótipo de acetilação de um determinado indivíduo. Entretanto, o grande número de variantes alélicas de *NAT2* dificulta a genotipagem por PCR-RFLP devido à grande quantidade de reações a serem realizadas para a caracterização completa de todos os alelos em uma única amostra, além de impossibilitar a identificação de novos polimorfismos. O sequenciamento direto do produto amplificado contendo a região codificante de *NAT2* permitiu uma maior discriminação dos alelos, possibilitando a correta classificação do perfil de acetilação dos indivíduos arrolados nesse estudo. Além disso, a seleção dos pacientes no presente estudo foi realizada de forma a excluir ao máximo variáveis que pudessem interferir em nossos resultados e assim contribuir para um melhor entendimento dos mesmos.

Nenhuma associação foi observada para os demais *loci* estudados (*CYP2E1*, *GSTM1* e *GSTT1*). Apesar das variantes gênicas apresentarem frequências similares àquelas encontradas por outros autores em nossa população (Sigimura *et al.*, 1995; Suarez-Kurtz *et al.*, 2007), não verificamos uma relação das mesmas com hepatotoxicidade ou hepatite medicamentosa induzida por drogas anti-TB (Tabelas 1 e 2 do Capítulo 4). Poucos são os estudos avaliando a influência desses genes na ocorrência de efeitos colaterais em pacientes com TB, e ainda assim apresentam resultados contraditórios (Roy *et al.*, 2001; Huang *et al.*, 2003; Cho *et al.*, 2007; Huang *et al.*, 2007; Leiro *et al.*, 2008). Estudos com um maior tamanho amostral envolvendo outras populações devem, pois ser realizados, para melhor elucidar o possível



envolvimento dos polimorfismos nos genes *CYP2E1*, *GSTM1* e *GSTT1*, na ocorrência de efeitos adversos induzidos por drogas anti-TB.

A possibilidade de avaliação da influência da variabilidade genética do hospedeiro, em relação a desfechos comuns em resposta ao tratamento anti-TB com esquemas contendo INH, é sem dúvida inovadora, e certamente contribuirá com informações importantes para a reavaliação da conduta terapêutica individualizada e um melhor manejo do tratamento anti-TB. Isso permitirá, a nível profilático, a substituição da INH por outra droga, para a prevenção do desenvolvimento de ADRs, em pacientes cujo perfil de acetilação seja predominantemente lento, ou ainda a correção da dosagem de INH nos pacientes com um predomínio de alelos funcionais de *NAT2*.





## *5- CONCLUSÕES*



- A fração da população brasileira, aqui representada por indivíduos residentes no Rio de Janeiro e em Goiás, apresenta uma grande diversidade de alelos de *NAT2*, resultado da miscigenação das populações parentais (europeus, africanos e ameríndios);
- O sete SNPs de *NAT2* mais frequentes no mundo foram identificados na população estudada assim como outros SNPs ainda não descritos na literatura, contribuindo para a diversidade alélica de *NAT2*;
- A fração da população brasileira aqui estudada apresentou uma quantidade maior de alelos de *NAT2* com baixa atividade e conseqüentemente um maior número de indivíduos com perfil de acetilação lenta em relação aos acetiladores intermediários e rápidos;
- Os pacientes com tuberculose, sob tratamento anti-TB, que tinham perfil de acetilação lenta apresentaram picos de transaminases significativamente maiores quando comparados àqueles indivíduos portadores de genótipos/fenótipos de acetilação rápida ou intermediária;
- Os pacientes com tuberculose, sob tratamento anti-TB com esquema contendo isoniazida, portadores de um genótipo/fenótipo de acetilação lenta apresentaram um risco aumentado em desenvolver hepatotoxicidade ou hepatite medicamentosa em relação aos acetiladores rápidos/intermediários;
- Nenhuma associação foi observada entre polimorfismos nos genes de *CYP2E1*, *GSTM1* e *GSTT1* com a ocorrência de efeitos adversos induzidos por drogas anti-TB, na população estudada do Rio de Janeiro.



## ***6- REFERÊNCIAS BIBLIOGRÁFICAS***



Agúndez JA, Ladero JM, Olivera M, Lozano L, Fernández-Arquero M, de laConcha EG, et al. N-acetyltransferase 2 polymorphism is not related to the risk of advanced alcoholic liver disease. *Scand J Gastroenterol.* 2002; 37:99-103.

Agúndez JAG. NAT2 Genotyping: Equilibrium between accuracy and feasibility in routine analyses. *J Appl Res.* 2003; 3:1-6.

Bagheri H, Michel F, Lapeyre-Mestre M, Lagier E, Cambus JP, Valdiguíé P, et al. Detection and incidence of drug-induced injuries in hospital: a prospective analysis from laboratory signal. *J Clin Pharmacol.* 2000; 50: 479-484.

Bell DA, Taylor JA, Buttler 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.

Bissell DM, Gores GJ, Laskin DL, Hoofnagle JH. Drug-induced liver injury: mechanisms and test systems. *Hepatology.* 2001; 33:1009-1013.

Blanchard JS. Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Annu Rev Biochem.* 1996; 65: 215-239.

Blum M, Demierre A, Grant DM, Heim M, Meyer UA. Molecular mechanism of slow acetylation of drugs and carcinogens in human. *Proc Natl Acad Sci.* 1991; 88:5237-5241.

Brockmöller J, Kerb R, Drakoulis N, Staffeldt B, Roots I. Glutathione S-transferase M1 and its variants A and B as host factors of bladder cancer susceptibility: a case-control study. *Cancer Res.* 1994; 54:4103-4111.

Caro AA, Cederbaum AI. Oxidative stress, toxicology, and pharmacology of CYP2E1. *Annu Rev Pharmacol Toxicol.* 2004; 44:27-42.

Carriere V, Berthou F, Baird S, Belloc C, Beaune P, Waziers ID. Human cytochrome P4502E1 (CYP2E1): from genotype to phenotype. *Pharmacogenetics.* 1996; 6:203-211.



Cascorbi I, Drakoulis N, Brockmüller 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.

Cavaco I, Reis R, Gil JP, Ribeiro V. *CYP3A4\*1B* and *NAT2\*14* alleles in a native African population. *Clin Chem Lab Med.* 2003; 41:606-609.

Cho HJ, Koh WJ, Ryu YJ, Ki CS, Nam MH, Kim JW, et al. Genetic polymorphisms of *NAT2* and *CYP2E1* associated with antituberculosis drug-induced hepatotoxicity in Korean patients with pulmonary tuberculosis. *Tuberculosis.* 2007; 87:551-556.

Cotran RS, Kumar V, Robbins SL, Schoen FJ, editores. *Robbins - Patologia Estrutural e Funcional*, 5ª ed. Rio de Janeiro: Guanabara Koogan, 1996.

Cotton SC, Sharp L, Little J, Brockton N. Glutathione S-transferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol.* 2000; 151:7-32.

Coughlin SS, Hall IJ. Glutathione S-transferase polymorphisms and risk of ovarian cancer: A HuGE review. *Genet Med.* 2002; 4:250–257.

Dalcomo M, Melo FAF, Pinto WP. II Consenso Brasileiro de Tuberculose: Diretrizes Brasileiras para Tuberculose. *J Bras Pneumol.* 2004, v. 30, suppl. 1. Brasília: LILACS, Periodica e SciELO, 2004.

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.

Das-Roy P, Majumder M, Roy B. Pharmacogenomics of anti-TB drugs-related hepatotoxicity. *Pharmacogenomics.* 2008; 9:311-321.



Deitz AC, Zheng W, Leff MA, Gross M, Wen WQ, Doll MA, et al. N-Acetyltransferase-2 genetic polymorphism, well-done meat intake, and breast cancer risk among postmenopausal women. *Cancer Epidemiol Biomarkers Prev.* 2000; 9:905-10.

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.

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.

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; 231:413-20

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.

Dupret JM, Grant DM. Site-directed mutagenesis of recombinant human arylamine N-acetyltransferase expressed in *Escherichia coli*. Evidence for direct involvement of Cys<sup>68</sup> in the catalytic mechanism of polymorphic human NAT2. *J Biol Chem.* 1992; 267: 7381-7385.

Evans DAP. N-acetyltransferase. *Pharmac Ther.* 1989; 42:157-234.

Flynn JL, Chan J. Tuberculosis: Latency and Reactivation. *Infect Immun.* 2001a; 69:4195-4201.

Flynn JL, Chan J. Immunology of tuberculosis. *Annu Rev Immunol.* 2001b; 19: 93-129.

Fretland AJ, Leff MA, Doll MA, Hein DW. Functional characterization of human N-acetyltransferase 2 (*NAT2*) single nucleotide polymorphisms. *Pharmacogenetics.* 2001; 11:207-215.



Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C. Tuberculosis. *Lancet*. 2003; 362:887-899.

Garibaldi RA, Drusin RE, Ferebee SH, Gregg MB. Isoniazid-associated hepatitis. Report of an outbreak. *Am Rev Respir Dis*. 1972; 106:357-365.

Geisler SA, Olshan AF. GSTM1, GSTT1, and the Risk of Squamous Cell Carcinoma of the Head and Neck: A Mini-HuGE Review. *Am J Epidemiol*. 2001; 154:95–105.

Grant DM. Structures of human arylamine N-acetyltransferases. *Curr Drug Metab*. 2008; 9:465-470.

Guttmacher, AE, Collins, FS. Genomic medicine – A primer. *N Engl J Med*. 2002; 347:1512-1520.

Harada S, Abei M, Tanaka N, Agarwal DP, Goedde HW. Liver glutathione S-transferase polymorphism in Japanese and its pharmacogenetic importance. *Hum Genet*. 1987; 75:322-325.

Hardman JG, Limberd LE, Molinoff PB, Ruddon RW, Gilman AG, editores. *Goodman & Gilman: as bases farmacológicas da terapêutica*. 9ª.ed. Rio de Janeiro: McGraw-Hill, 1996.

Harrison TR, Braunwld E, Fauci AS, Kasper DL, Hauser SL, editores. *Distúrbios do sistema gastrointestinal*. Medicina Interna. Rio de Janeiro: McGraw-Hill, 2002.

Hayes JD, Flanagan JU, Jowsey IR. Glutathione Ttransferases. *Annu Rev Pharmacol Toxicol*. 2005; 45:51-88.

Hein DW, Doll MA, Fletland 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.



Hein DW. Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. *Mut Res.* 2002; 506-507:65-77.

Henrion-Caude A, Roussey C, Housset M, Flahault C, Fryer A, Chadelat AA, et al. Liver disease in pediatric patients with cystic fibrosis is associated with glutathione S-transferase P1 polymorphism. *Hepatology.* 2002; 36:913-917.

Higuchi N, Tahara N, Yanagihara K, Fukushima H, 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.

Hickman D & Sim E. N-acetyltransferase polymorphism. Comparison of phenotype and genotype in humans. *Biochem Pharmacol.* 1991; 42:1007-1014.

Hiratsuka, Kishikawa Y, Takekuma Y, Matsuura M, Narahara K, Inoue T, et al. Genotyping of the N-acetyltransferase 2 polymorphism in the prediction of adverse drug reactions to isoniazid in Japanese patients. *Drug Metabol Pharmacokin.* 2002; 17:357-362.

Holley SL, Fryer AA, Haycock JW, Grubb SEW, Strange RC, Hoban PR. Differential effects of glutathione S-transferase pi (GSTP1) haplotypes on cell proliferation and apoptosis. *Carcinogenesis.* 2007; 28:2268-2273.

Holton SJ, Dairou J, Sandy J, Rodrigues-Lima F, Dupret JM, Noble ME, et al. Structure of *Mesorhizobium loti* arylamine N-acetyltransferase 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2005; 61:14-16.

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.

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.





Huang YS. Genetic polymorphisms of drug-metabolizing enzymes and the susceptibility to antituberculosis drug-induced liver injury. *Expert Opin Drug Metab Toxicol.* 2007; 3:1-8.

Huang YS, Su WJ, Huang YH, Chen CY, Chang FY, Lin HC, et al. Genetic polymorphisms of manganese superoxide dismutase, NAD(P)H:quinine oxidoreductase, glutathione S-transferase M1 and T1, and the susceptibility to drug-induced liver injury. *J Hepatol.* 2007; 47:128-134.

Hug H, Bagatto D, Dannecker R, Schindler R, Horlancher O, Gut J. ADRIS – The adverse drug reactions information scheme. *Pharmacogenetics.* 2003; 13:767-772.

IBGE – Instituto Brasileiro de Geografia Estatística. Brasil: 500 Anos de Povoamento. Rio de Janeiro: IBGE, 2000.

Ingelman-Sundberg M. Pharmacogenetics: an opportunity for a safer and more efficient pharmacotherapy. *J Intern Med.* 2001; 250: 186-200.

Ingelman-Sundberg M. Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol Sci.* 2004; 25:193-200.

Jorge-Nebert LF, Eichelbaum M, Griese EU, Inaba T, Arias T.D. 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.

Kalow W. Pharmacogenetics in Biological Perspective. *Pharmacol Rev.* 1997; 49:369-379.

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.

Kirchheiner J, Fuhr U, Brockmöller J. Pharmacogenetics-based therapeutic recommendations-ready for clinical practice? *Nat Rev Drug Discov.* 2005; 4:639-47.



Kita T, Tanigawara Y, Chikazawa S, Hatanaka H, Sakaeda T, Komada F, et al. N-acetyltransferase 2 genotype correlated with isoniazid acetylation in Japanese tuberculous patients. *Biol Pharm Bull.* 2001; 24: 544-549.

Knight RA, Selin MJ, Harris HW. Genetic factors influencing isoniazid blood levels in humans. *Trans. 18<sup>th</sup> Conf. Chemother. Tuberc. Washinton D. C., Veterans Adm. Army Navy:* 1959; 52-58.

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.

Landi S. Mammalian class theta GST and differential susceptibility to carcinogens: a review. *Mut Res.* 2000; 463:247-283.

Larrey D. Drug-induced liver diseases. *J Hepatol.* 2000; 32: 77-88.

Larrey D. Epidemiology and individual susceptibility to adverse drug reactions affecting the liver. *Semin Liver Dis.* 2002; 22:145-155.

Lee WM. Medical progress: drug-induced hepatotoxicity. *N Engl J Med.* 2003; 349:474-485.

Leiro V, Fernandez-Villar A, Valverde D, Constenla L, Vazquez R, Pineiro L, et al. Influence of glutathione S-transferase M1 and T1 homozygous null mutations on the risk of antituberculosis drug-induced hepatotoxicity in Caucasian population. *Liver Int.* 2008; 28:835-839.

Lin HJ, Han CH, 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.

Lundkvist J, Jonsson B. Pharmacoeconomics of adverse drug reactions. *Fundam Clin Pharmacol.* 2004; 18: 275-280.



Mashimo M, Suzuki T, Abe M, Deguchi T. Molecular genotyping of N-acetylation polymorphism to predict phenotype. *Hum Genet.* 1992; 90:139-43.

McLellan RA, Oscarson M, Alexandrie AK, Seidegard J, Evans DA, Rannug A, et al. Characterization of a human glutathione S-transferase mu cluster containing a duplicated *GSTM1* gene that causes ultrarapid enzyme activity. *Mol Pharmacol.* 1997; 52:958-965.

Meyer UA, Zanger UM. Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annu Rev Pharmacol Toxicol.* 1997; 37:269-96.

Middlebrook G. Sterilization of tubercle bacilli by isonicotinic acid hydrazide and the incidence of variants resistant to the drug in vitro. *Am Rev Tuberc.* 1952; 65:765-767.

Mitchell JR, Thorgeisson UP, Black M, Timbrell JA, Snodgrass WR, Potter WZ, et al. Increased incidence of isoniazid hepatitis in rapid acetylators: possible relation to hydralazine metabolites. *Clin Pharmacol Ther.* 1975; 18:70-79.

Mitchell JR, Snodgrass WR, Gillette JR. The Role of Biotransformation in Chemical-Induced Liver Injury. *Environ Health Perspect.* 1976; 15: 27-38.

Mittal RD, Srivastava DS, Mandhani A. *NAT2* gene polymorphism in bladder cancer: a study from North India. *Int Braz J Urol.* 2004; 30:279-85.

Moussa LA, Khassouani CE, Hue B, Jana M, Bégau, B, Soulayman R. Determination of the acetylato phenotype in Moroccan tuberculosis patients using isoniazida as metabolic probe. *Int J Clin Pharmacol Ther.* 2002; 40:548-553.

Mrozikiewicz PM, Cascorbi I, Brockmöller J, Roots I. Determination and allelic allocation of seven nucleotide transitions within the Arylamine N-acetyltransferase gene in the Polish population. *Clin Pharmacol Ther.* 1996; 59:376-382.

Nebert DW, Russell DW. Clinical importance of the cytochromes P450. *Lancet.* 2002; 360:1155-1162.



Nelson SD, Mitchell JR, Timbrell JA, Snodgrass WR, Corcoran GB. Isoniazid and iproniazid: activation of metabolites to toxic intermediates in man and rat. *Science*. 1976; 193(4256):901-903.

Neuhaus, T, Ko, Yon-Dschun, Lorenzen K, Fronhoffs S, Harth V, et al. Association of cytochrome P450 2E1 polymorphisms and head and neck squamous cell cancer. *Toxicology*. 2004; 151:273-282.

Ohno M, Yamaguchi I, Yamamoto I, Fukuda T, Yolota 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.

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.

Parl FF. Glutathione S-transferase genotypes and cancer risk. *Cancer Lett*. 2005; 221:123-129.

Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SDJ. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci*. 2003; 100:177-182.

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*. 2006a; 78:423-436.

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*. 2006b; 27:720-730.

Pelkonen O, Turpeinen M, Hakkola J, Honkakoski P, Hukkanen J, Raunio H. Inhibition and induction of human cytochrome P450 enzymes: current status. *Arch Toxicol*. 2008; 82:667-715.



Pemble S, Schroeder KR, Spencer SR, Meyer DJ, Hallier E, Bolt HM, et al. Human glutathione S-transferase Theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J.* 1994; 300:271-276.

Pompeo F, Brooke E, Kawamura A, Mushtaq A, Sim E. The pharmacogenetics of NAT: structural aspects. *Pharmacogenomics.* 2002; 3:19-30.

Powell H, Kitteringham NR, Pirmohamed M, Smith DA, Perk BK. Expression of cytochrome P-4502E1 in human liver: assessment by mRNA, genotype and phenotype. *Pharmacogenetics.* 1998; 8:411-421.

Prows CA, Prows DR. Medication Selection by genotype: How genetics is changing drug prescribing and efficacy. *Am J Nurs.* 2004; 104:60-70.

Raimondi S, Paracchini V, Autrup H, Barros-Dios JM, Benhamou S, Boffetta P, et al. Meta-and Pooled Analysis of GSTT1 and Lung Cancer: A HuGE-GSEC Review. *Am J Epidemiol.* 2006; 164:1027–1042

Rastogi N & David HL. Mode of action of antituberculous drugs and mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Res Microbiol.* 1993; 144:133-143.

Rodrigues-Lima F, Dupret JM. 3D model of human arylamine N-acetyltransferase 2: structural basis of the slow acetylator phenotype of the R64Q variant and analysis of the active-site loop. *Biochem Biophys Res Commun.* 2002; 291:116-123.

Roses AD. Pharmacogenetics place in modern medical science and practice. *Life Sci.* 2002; 70:1471-1480.

Roy B, Chowdhury A, Kundu S, Santra A, Dey B, Chakraborty M, et al. Increased risk of antituberculosis drug-induced hepatotoxicity in individuals with glutathione S-transferase M1 “null” mutation. *J Gastroenterol Hepatol.* 2001; 16:1033-1037.



Roy B, Ghosh SK, Sutraghar D, Sikdar N, Mazumder S, Barman S. Predisposition of antituberculosis drug induced hepatotoxicity by cytochrome P450 2E1 genotype and haplotype in pediatric patients. *J Gastroenterol Hepatol*. 2006; 21:781-786.

Rozwarski DA, Grant GA, Barton DH, Jacobs-Jr WR, Sacchettini JC. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* 1998; 279: 98-102.

Sandy J, Mushtaq A, Kawamura A, Sinclair J, Sim E, Noble M. The structure of arylamine N-acetyltransferase from *Mycobacterium smegmatis* - an enzyme which inactivates the anti-tubercular drug, isoniazid. *J Mol Biol*. 2002; 318:1071-1083.

Santos-Filho ET. *Public Health Watch*. Políticas de TB no Brasil - Uma perspectiva da Sociedade Civil. Rio de Janeiro: *Open Society Institute*, 2006.

Sarma GR, Immanuel C, Kailasam S, Narayana ASL, Venkatesan P. Rifampin-induced release of hydrazine from isoniazid: a possible cause of hepatitis during treatment of tuberculosis with regimens containing isoniazid and rifampin. *Am Rev Respir Dis*. 1986; 133:1072-1075.

Schluger NW. Recent advances in our understanding of human host responses to tuberculosis. *Respir Res*. 2001; 2: 157-167.

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 *LICAM* in the Japanese population. *J Hum Genet*. 2001; 46:314-319.

Shimiku Y, Dobashi K, Mita Y, Endou K, Moriya S, Osano K, et al. DNA microarray genotyping of N-acetyltransferase 2 polymorphism using carbodiimide as the linker for assessment of isoniazid hepatotoxicity. *Tuberculosis*. 2006; 86:374-381.

Sigimura H, Hamada GS, Suzuki I, Iwase T, Kiyokawa E, Kino I, et al. CYP1A1 and CYP2E1 polymorphism and lung cancer, case-control study in Rio de Janeiro, Brazil. *Pharmacogenetics*. 1995; 5:S145-148.



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; 9:2435-41.

Sinclair JC, Sandy J, Delgoda R, Sim E, Noble MEM. Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat Struct Biol.* 2000; 7:560-564.

Singh J, Arora A, Garg PK, Thakur VS, Pande JN, Tandon RK. Antituberculosis treatment-induced hepatotoxicity: role of predictive factors. *Postgrad Med J.* 1995a; 71:359-362.

Singh J, Garg PK, Thakur VS, Tandon RK. Antitubercular treatment induced hepatotoxicity: does acetylator status matter? *Indian J Physiol Pharmacol.* 1995b; 39:43-46.

Slayden RA & Barry CE. The genetics and biochemistry of isoniazid resistance in *Mycobacterium tuberculosis*. *Microbes Infect.* 2000; 2:659-669.

Sodhi CP, Rana SV, Mehta SK, Vaiphei K, Attri S, Thakur S, et al. Study of oxidative stress in isoniazid-induced hepatic injury in young rats with and without protein-energy malnutrition. *J Biochem Toxicol.* 1996; 11: 139-146.

Strange RC, Jones PW, Fryer AA. Glutathione S-transferase: genetics and role in toxicology. *Toxicol Lett.* 2000; 112-113:357-363.

Strange RC, Spiteri MA, Ramachandran S, Fryer AA. Glutathione-S-transferase family of enzymes. *Mut Res.* 2001; 482:21-26

Suarez-Kurtz G, Vargens DD, Struchiner, CJ, Bastos-Rodrigues L, Pena SDJ. Self-reported skin color, genomic ancestry and the distribution of *GST* polymorphisms. *Pharmacogenet Genomics.* 2007; 17:765-771.

Sunahara S, Urano M, Ogawa M. Genetical and geographic studies on isoniazid inactivation. *Science.* 1961;134:1530-1531.



Takayama K, Wang L, David HL. Effect of Isoniazid on the In Vivo Mycolic Acid Synthesis, Cell Growth, and Viability of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 1972; 2: 29–35.

Tufariello JM, Chan J, Flynn JL. Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *Lancet*. 2003; 3: 578-590.

van Crevel R, Ottenhoff THM, van der Meer JWM. Innate immunity to *Mycobacterium tuberculosis*. *Clin Microbiol Rev*. 2002; 15: 294-309.

Vineis P, Veglia F, Anttila S, Benhamou S, Clapper ML, Dolzan V, et al. *CYP1A1*, *GSTM1* and *GSTT1* polymorphisms and lung cancer: a pooled analysis of gene-gene interactions. *Biomarkers*. 2004; 9:298-305.

Vuilleumier N, Rossier MF, Chiappe A, Degoumois F, Dayer P, Mermillod B., et al. CYP2E1 genotype and isoniazid-induced hepatotoxicity in patients treated for latent tuberculosis. *Eur J Clin Pharmacol*. 2006; 62:423-429.

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.

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.

Wang H, Vath GM, Gleason KJ, Hanna PE, Wagner CR. Probing the mechanism of hamster arylamine N-acetyltransferase 2 acetylation by active site modification, site-directed mutagenesis, and pre-steady state and steady state kinetic studies. *Biochemistry*. 2004; 43:8234-8246.

Watanabe M, Sofuni T, Nohmi T. Involvement of Cys69 residue in the catalytic mechanism of N-hydroxyarylamine O-acetyltransferase of *Salmonella typhimurium*. Sequence similarity at the amino acid level suggests a common catalytic mechanism of acetyltransferase for *S. typhimurium* and higher organisms. *J Biol Chem*. 1992; 267:8429-8436.





Watanabe J, Hayashi S, Kawajiri K. Different regulations and expression of the human CYP2E1 gene due to the Rsa I polymorphism in the 5`flanking region. *J. Biochem.* 1994; 116:321-326.

Westwood IM, Holton SJ, Rodrigues-Lima F, Dupret JM, Bhakta S, Noble MEM, et al. Expression, purification, characterization and structure of *Pseudomonas aeruginosa* arylamine N-acetyltransferase. *Biochem J.* 2005; 385:605-612.

Westwood IM, Kawamura A, Fullam E, Russell AJ, Davies SG, Sim E. Structure and mechanism of arylamine N-acetyltransferases. *Curr Top Med Chem.* 2006; 6:1641-1654.

Wijnen PAHM, Op Den Buijsch RAM, Drent M, Kuipers PMJC, Neef C, Bast A, et al. Review article: the prevalence and clinical relevance of cytochrome P450 polymorphisms. *Aliment Pharmacol Ther.* 2007; 26 Supl:211-219.

Widersten M, Pearson WR, Engström A, Mannervik B. Heterologous expression of the allelic variant mu-class glutathione transferases mu and psi. *Biochem J.* 1991; 276:519-524.

Winder FG, Collins PB. Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J Gen Microbiol.* 1970; 63:41-48.

Winder FG, Collins PB, Rooney SA. Effects of isoniazid on mycolic acid synthesis in *Mycobacterium tuberculosis* and on its cell envelope. *Biochem J.* 1970; 117:127P.

Woodward KN, Timbrell JA. Acetylhydrazine hepatotoxicity: the role of covalent binding. *Toxicology.* 1984; 30: 65-74.

World Health Organization. Global tuberculosis control. WHO report 2000. Geneva, Switzerland, WHO/CDS/TB/2000.275.

World Health Organization. Global tuberculosis control: surveillance, planning, financing. WHO report 2008. Geneva, Switzerland, WHO/HTM/TB/2008.393.



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.

Yokogawa K, Nakaharu T, Ishizaki J, Ozaki E, Takeda Y, Mabuchi H, et al. Kinetic Phenotypic diagnosis of N-acetylation polymorphism in patients based on ratio of urinary metabolites of salicylazosulfapyridine *Int J. Pharmac.* 2001; 229:183-191

Zang Y, Doll MA, Zhao S, States JC, Hein DW. Functional characterization of single polymorphisms and haplotypes of human N-acetyltransferase 2. *Carcinogenesis.* 2007; 28:1665-1671.

Zhang Y & Young DB. Molecular mechanisms of isoniazid: a drug at the front line of tuberculosis control. *Trens Microbiol.* 1993; 1:109-113.

**Sites:**

[www.louisville.edu/medschool/pharmacology/NAT.html](http://www.louisville.edu/medschool/pharmacology/NAT.html) (visitado em 17-03-09)

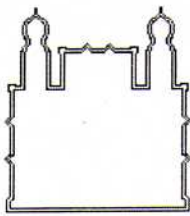
[www.cypalleles.ki.se/cyp2e1.htm](http://www.cypalleles.ki.se/cyp2e1.htm) (visitado em 17-03-09)



## ***7- ANEXOS***



## *ANEXO 1*



Ministério da Saúde  
Fundação Oswaldo Cruz  
COMITÊ DE ÉTICA EM PESQUISA-CEP/FIOCRUZ

Rio de Janeiro, 16 de setembro de 2003.

Carta: 098/03


De: CEP/FIOCRUZ

Para: - Dr. Adalberto Rezende Santos e  
- Dra. Marli Maria Lima

Prezados Senhores,

Estamos encaminhando o parecer do protocolo **21603**, intitulado: **“Estudo de associação entre polimorfismos de base única no gene de NAT2 e a presença de reações adversas em pacientes com Tuberculose”** que está **APROVADO**.

Atenciosamente

  
CARLA DIAS NETTO  
Secretária Geral  
Comitê de Ética em Pesquisa  
Fundação Oswaldo Cruz



## *ANEXO 2*



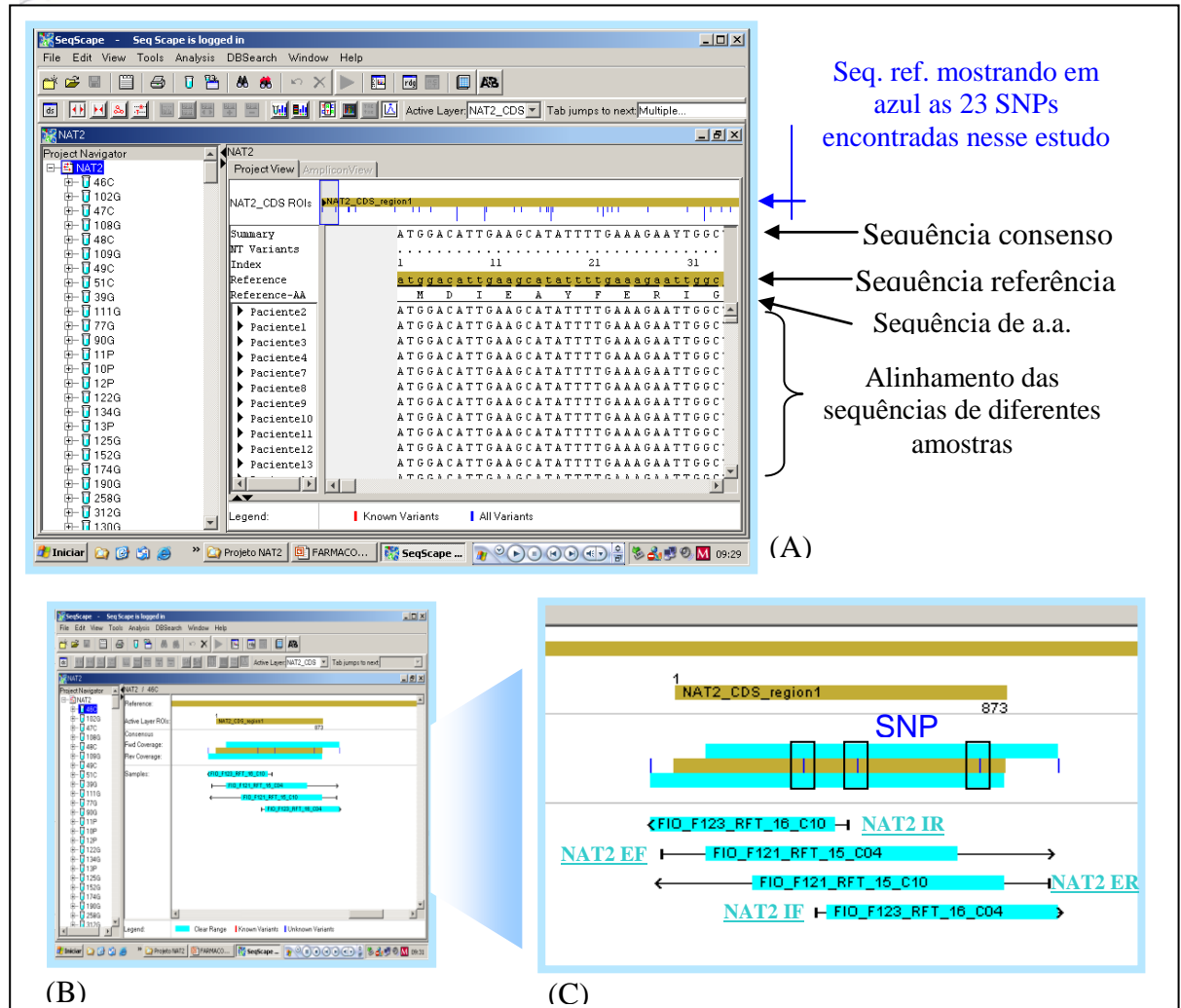


Figura 29: Ilustração da utilização do software SeqScape v. 2.5. (a) Nessa janela pode ser observado o alinhamento das amostras bem como a sequência referência, sequência consenso e a sequência de aminoácidos de NAT2. (b e c) Nessas janelas podem ser observadas as sequências obtidas a partir de cada iniciador bem como a região de aproveitamento de cada uma delas (azul claro). Nesse indivíduo foram detectados 3 SNPs, ou seja, três bases se apresentaram distintas da sequência referência (c).

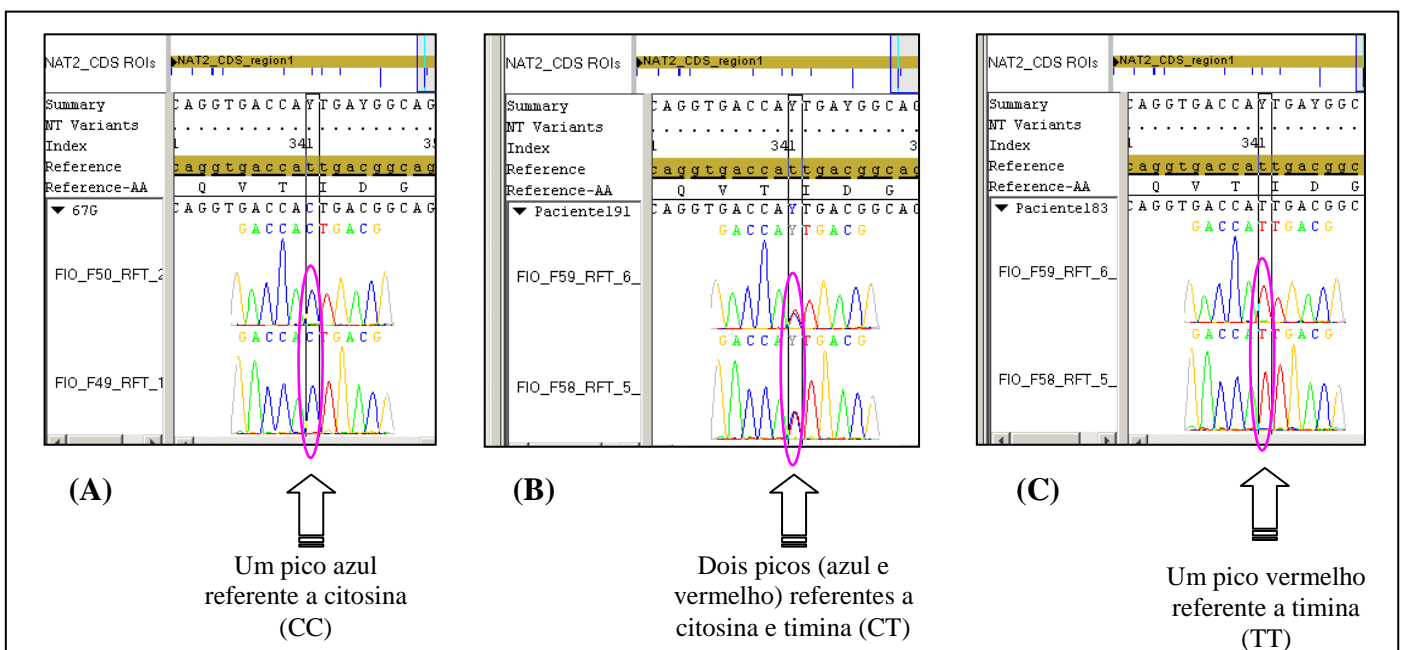
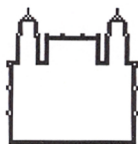


Figura 30: Identificação de SNP através da visualização do cromatograma dentro do programa SeqScape. A figura mostra três amostras diferentes e os possíveis genótipos para o SNP T341C. (A) Amostra mutante homocigota, (B) amostra mutante heterocigota e (C) amostra tipo-selvagem para o SNP T341C.





## *ANEXO 3*



Ministério da Saúde

FIOCRUZ  
Fundação Oswaldo Cruz

Presidência

## DECLARAÇÃO

Declaro, para os devidos fins, que **Raquel Lima de Figueiredo Teixeira** figura no rol de inventores do pedido de patente intitulado:

“Método para a predição do desfecho do tratamento de doenças humanas utilizando fármacos metabolizados pela N-acetiltransferase 2 humana (NAT2) com base em polimorfismos genéticos”,

sob o nº PI0704093-8, depositado em 04 de outubro de 2007, no Instituto Nacional da Propriedade Industrial. O presente pedido de patente atua como prioritário para o depósito de correspondente internacional na Organização Mundial da Propriedade Industrial (OMPI), através do Tratado de Cooperação em Matéria de Patentes (Patent Cooperation Treaty – PCT), sob o nº PCT/BR2008/000304, realizado em 06 de outubro de 2008.

Rio de Janeiro, 31 de março de 2009

Leila Longa  
Coordenação de Gestão Tecnológica  
Vice-Presidência de Produção e Inovação em Saúde  
Presidência/FIOCRUZ



## *ANEXO 4*



**ESTRATÉGIA DE CLONAGEM PARA RE-SEQUENCIAMENTO DE NAT2 HUMANO**

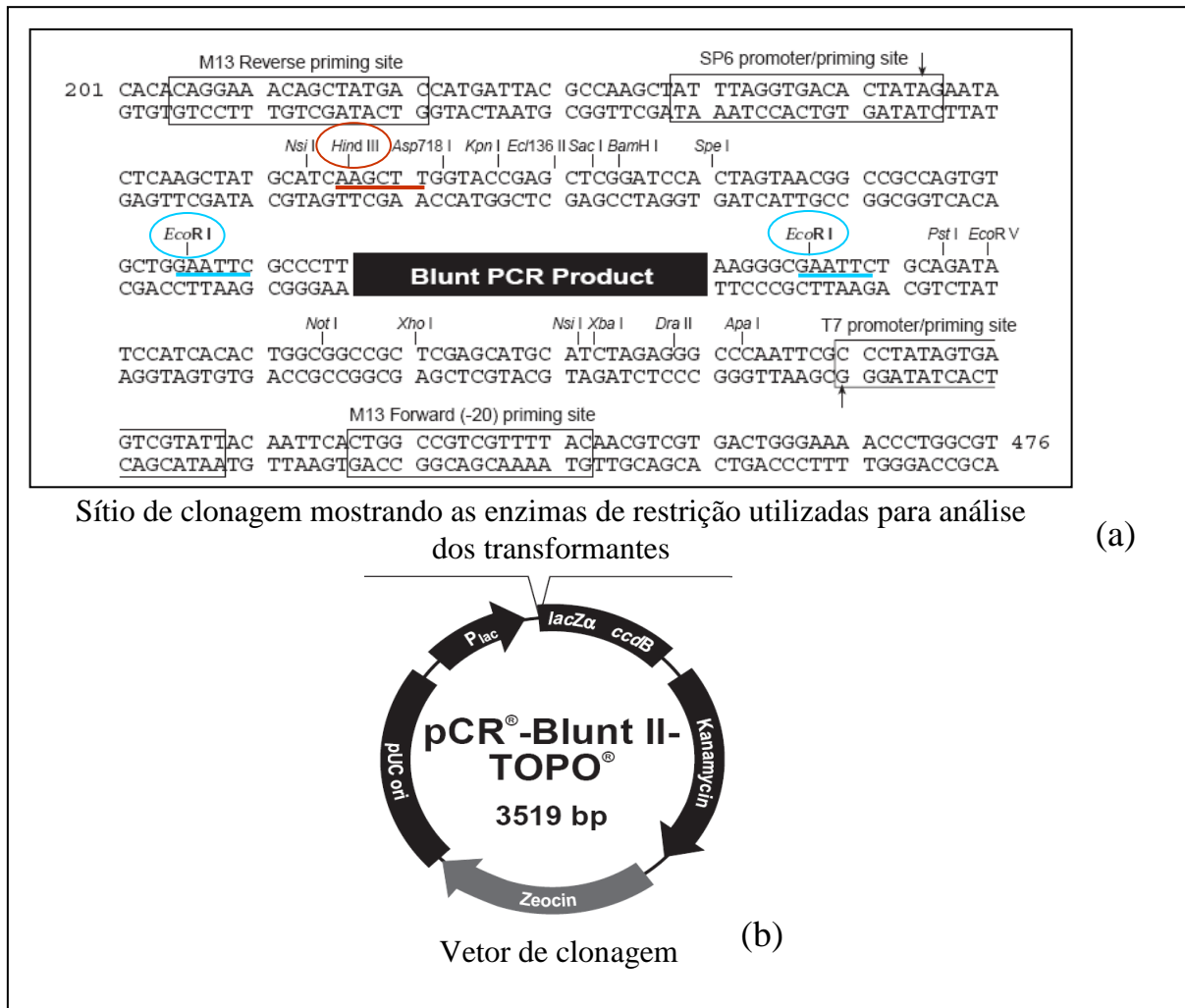


Figura 31: (a) Sequência nucleotídica da região de clonagem do vetor pCR-BluntII-TOPO indicando os sítios de restrição de *Hind* III (apenas um sítio no vetor) e *Eco* RI (dois sítios que flanqueiam o inserto). (b) Esquema do vetor de clonagem indicando a localização da entrada do inserto dentro da região codificante do gene “suicida” (*ccdB*).

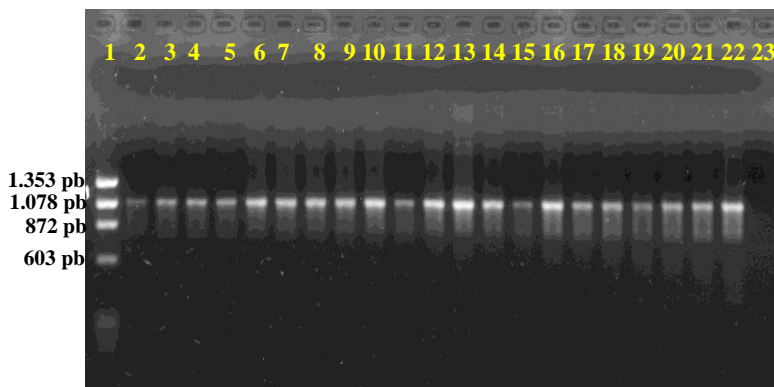


Figura 32: Gel de agarose 1,5% com PCR *NAT2* a partir de colônia para identificação de bactérias portadoras de pCRBlunt-II/*NAT2*. (1) DNA  $\phi$ X174 digerido com *Hae* III (2) produto de PCR (1093 pb) gerado a partir de uma amostra de DNA humano (controle positivo) (3 a 22) produtos de PCR derivados de 20 diferentes transformantes; (23) controle negativo.

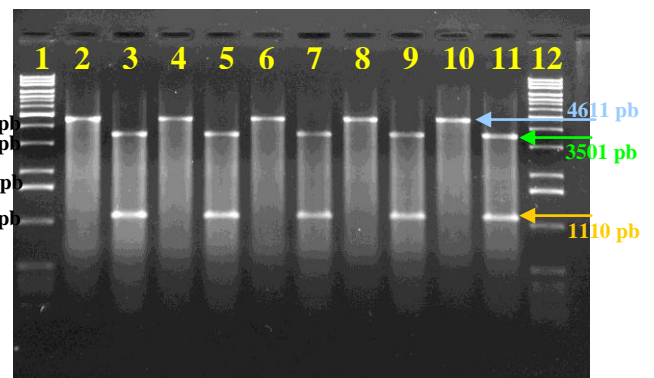


Figura 33: Análise de restrição para confirmação da presença do inserto dos clones pCRBlunt-II/*NAT2*. (1 e 12) 1Kb DNA Extension Ladder; (2, 4, 6, 8 e 10) digestão dos clones com *Hind* III; (3, 5, 7, 9 e 11) digestão dos clones com *Eco* RI.



## *ANEXO 5*

## Parecer Consubstanciado de Projeto de Pesquisa

**Título do Projeto: Influência da variabilidade genética humana dos genes NAT2, CYP2E1 e GST na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes com tuberculose ativa**

**Pesquisador Responsável Adalberto Rezende Soares**

**Data da Versão 24/09/2007**

**Cadastro 421/07**

**Data do Parecer 12/11/2007**

**Grupo e Área Temática I.1 Genética Humana**

### Objetivos do Projeto

Avaliar a possível associação entre variantes alélicas dos genes NAT2, CYP2E1 e GST na ocorrência de reações adversas hepáticas induzidas pela isoniazida em pacientes com tuberculose ativa e, HUCFF-UFRJ, RJ, referência para AIDS  
- det haplotipos de NAT2, frequência dos alelos c1 e c2 de CYP2E1, e de GSTM1\*0 e GSTT1\*0, correlacionar com hepatotoxicidade

### Sumário do Projeto

Tratamento atual da TB = isoniazida (INH) (e rifampicina (RIF) e pirazinamida (PZA)). Reações adversas a nível hepático ocorrem, envolvendo as enzimas N-acetiltransferase 2 (NAT 2) e a monooxigenase (CYP 450 2E1) e, possivelmente, glutationa S transferase (GST). O estudo pretende avaliar a associação entre variabilidade genética dessas enzimas e hepatotoxicidade da droga.

Itens Metodológicos e Éticos	Situação
Título	Adequado
Autores	Adequados
Local de Origem na Instituição	Adequado
Projeto elaborado por patrocinador	Não
Aprovação no país de origem	Não necessita
Local de Realização	Outro (citar no comentário)
Outras instituições envolvidas	Sim
Condições para realização	Adequadas

Comentários sobre os itens de Identificação

HUCFF-UFRJ + IOC-FIOCRUZ

Introdução	Adequada
------------	----------

Comentários sobre a Introdução

Objetivos	Adequados
-----------	-----------

Comentários sobre os Objetivos

Pacientes e Métodos	
Delineamento	Adequado
Tamanho de amostra	Total 245 Local RJ
Cálculo do tamanho da amostra	Não necessário (pesquisa qualitativa)
Participantes pertencentes a grupos especiais	Não
Seleção equitativa dos indivíduos participantes	Adequada
Critérios de inclusão e exclusão	Adequados
Relação risco- benefício	Adequada
Uso de placebo	Não utiliza
Período de suspensão de uso de drogas (wash out)	Não utiliza
Monitoramento da segurança e dados	Comentário
Avaliação dos dados	Adequada - qualitativa
Privacidade e confidencialidade	Adequada
Termo de Consentimento	Comentário
Adequação às Normas e Diretrizes	Sim

Comentários sobre os itens de Pacientes e Métodos  
Como usa material de pacientes em tratamento, o tratamento será acompanhado pelo Hospital normalmente

Cronograma	Adequado
Data de início prevista	novembro 2007
Data de término prevista	nov 09
Orçamento	Comentário
Fonte de financiamento externa	Não

Comentários sobre o Cronograma e o Orçamento

POM - IOCe orçamento regular HUCFF-UFRJ

Referências Bibliográficas	Adequadas
----------------------------	-----------

Comentários sobre as Referências Bibliográficas

Recomendação

**Aprovar**

Comentários Gerais sobre o Projeto

Um projeto simples, com objetivos bem escritos, que irá fazer parte de uma rotina bem estabelecida (HUCFF) e de tecnologias usadas rotineiramente nos laboratórios da FIOCRUZ, com bastante chance de trazer informações de grande utilidade para o tratamento de indivíduos infectados com M. tuberculosis no futuro.

Após análise por este colegiado, tendo por referência as normas e diretrizes da Resolução 196/96, foi decidido pela APROVAÇÃO do referido protocolo.

Solicitamos as seguintes correções:

- 1) Já que não há previsão para inclusão de menores de idade ou indivíduos que necessitarão de representante legal, retirar o parênteses "ou responsável legal" do Termo de Consentimento Livre e Esclarecido;
- 2) Solicitamos a correção da Folha de Rosto no relativo a Área de conhecimento. Está corretamente indicado como GENÉTICA, mas ao classificar o Grupo foi colocado como Grupo 2, quando o correto e Grupo 1.

Informamos, outrossim, que deverão ser apresentados relatórios parciais/anuais e relatório final do projeto de pesquisa.

Além disso, qualquer modificação ou emenda ao protocolo original deverá ser submetida para apreciação do CEP/Fiocruz.

**Marlene Braz**

Coordenadora do Comitê de Ética em Pesquisa  
Em Seres Humanos da Fundação Oswaldo Cruz



## *ANEXO 6*





## ESTRATÉGIA DE GENOTIPAGEM DE *CYP2E1*, *GSTM1* E *GSTT1* HUMANOS



Figura 34: Gel de agarose 2% com o perfil de digestão (*Rsa* I) dos três genótipos de *CYP2E1* para o SNP -1053 C→T (Ver Capítulo 4). (1 e 13) Marcador de peso molecular de 100 pb; (2) genótipo mutante homocigoto (c2/c2); (3 a 8) genótipo tipo-selvagem homocigoto (c1/c1); (9 a 12) genótipo

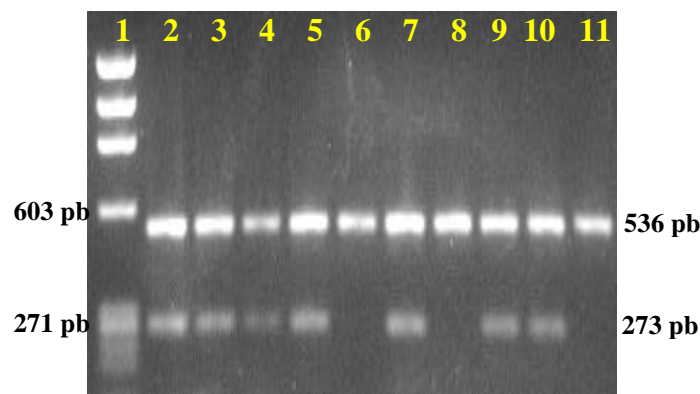


Figura 35: Gel de agarose 2,5% com PCR padrão para detectar a presença do gene *GSTM1* tipo-selvagem (*GSTM1*\*A ou *GSTM1*\*B) (Ver Capítulo 4). (1) DNA  $\phi$ X174 digerido com *Hae* III; (2, 3, 4, 5, 7, 9, 10) amostras *GSTM1*\*A-B/*GSTM1*\*A-B ou *GSTM1*\*A-B/*GSTM1*\*0; (6 e 8) amostras *GSTM1*\*0/*GSTM1*\*0.

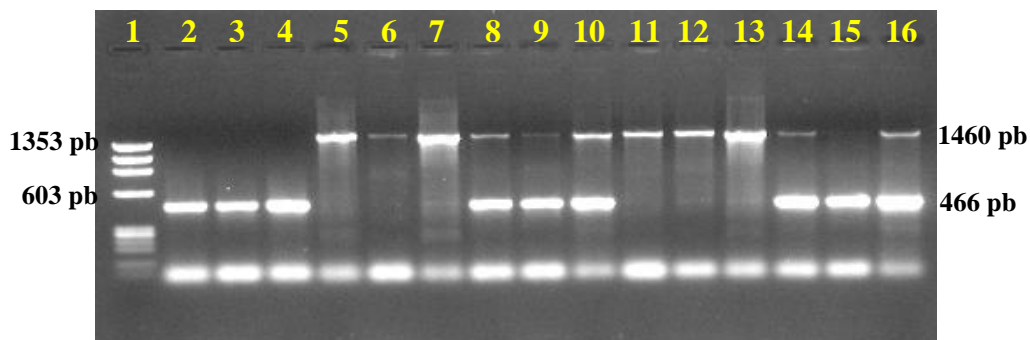


Figura 36: Gel de agarose 1,5% com PCR/*GSTT1* (Ver Capítulo 4): (1) DNA  $\phi$ X174 digerido com *Hae* III; (2, 3, 4, 15) amostras com genótipo *GSTT1*\*A/*GSTT1*\*A; (5, 6, 7, 11, 12, 13) amostras com genótipo *GSTT1*\*0/*GSTT1*\*0; (8, 9, 10, 14, 16) amostras com genótipo *GSTT1*\*A/*GSTT1*\*0.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)