Biochemical Pharmacology, Vol. 55, pp. 361–366, 1998.

ISSN 0006-2952/98/$19.00 + 0.00
PII S0006-2952(97)00478-4

Genotyping Human Arylamine *N*-Acetyltransferase Type 1 (NAT1)

The Identification of Two Novel Allelic Variants

Mark A. Payton and Edith Sim

Department of Pharmacology, University of Oxford, Mansfield Road, Oxford, OX1 3QT, UK

ABSTRACT. Human arylamine *N*-acetyltransferase (NAT) is known to exist as two isoenzymes, NAT1 and NAT2, with different though overlapping substrate specificities. NAT1 and NAT2 are polymorphic at both genetic and phenotypic levels with four distinct alleles described in Caucasians for NAT1. Though clear genotype/phenotype associations exist for NAT2, the same remains unclear for NAT1. Whole blood taken from 32 individuals were NAT1 genotyped and compared to previously obtained NAT1 activities using *p*-aminobenzoic acid as a substrate [1]. The NAT1 alleles of one individual, who had low NAT1 activity, were sequenced and compared to the wild type allele *NAT1*4*. A novel, non-conservative, substitution was present in both alleles at nucleotide position 560 and results in the exchange of an arginine for a glutamine at amino acid position 187. A glutamine is found in NAT2 at amino acid position 187 and has been implicated in substrate binding. This report describes a simple and effective genotyping method which detects the four previously reported NAT1 polymorphisms, and the described novel low acetylating polymorphism, by either NAT1 allele specific-PCR amplification or restriction fragment length polymorphism analysis of PCR amplified products. We suggest that NAT1 genotype/phenotype correlations will become more clear as further allelic variants are determined. BIOCHEM PHARMACOL **55**;3:361–366, 1998. © 1998 Elsevier Science Inc.

KEY WORDS. arylamine; acetyltransferase; NAT; genotype; phenotype; PCR-RFLP

Exposure to certain xenobiotics is associated with increased risk to the development of cancers, which include bladder and colorectal cancer [2–4]. Bladder cancer is the fourth most common cancer affecting males of UK origin [5]. Its development has been strongly linked to environmental and occupational exposures to chemicals [6, 7]. Such carcinogenic chemicals include those previously found in the dye and rubber manufacturing industries, and components of diesel fumes and cigarette smoke [8, 9]. A class of highly mutagenic compounds, the aromatic amines (or their metabolites), are found in these xenobiotic sources and have been linked to DNA-adduct formation and cancer development [10].

Aromatic amines and hydrazines are metabolised by arylamine *N*-acetyltransferase (NAT)† which catalyses the transfer of an acetyl group from acetyl CoA to the nitrogen of the substrate [11, 12]. In humans, NAT is actively expressed as two isoenzymes with distinct, but overlapping, substrate specificity. These isoenzymes, termed NAT1 and NAT2, have been mapped to the short arm of chromosome 8 [13] at position 8p22 within 400 kb of each other [14]. Tissue distribution of NAT1 is almost ubiquitous, in contrast, NAT2 has a rather more limited distribution found in the liver and epithelial cells of the intestine [15, 16]. The human NAT2 locus has demonstrated a high degree of allelic polymorphism, with certain genotypes significantly associated with bladder and breast cancer, particularly in cases of environmental exposure to arylamines [7, 9, 17]. Variations in the NAT2 genotype have been associated with differing degrees of acetylation rates [17, 18]. NAT1 has also been shown to exhibit allelic polymorphism [19].

The previously devised nomenclature for known NAT1 allelic variations has been based on an order of discovery, irrespective of the species of origin [19]. Thus far, four distinct allelic polymorphisms at the NAT1 locus within the human Caucasian population have been described and termed: *NAT1*3*, *NAT1*4*, *NAT1*10* and *NAT1*11* (Fig. 1). The allele *NAT1*4* is considered the wild type with a coding region of 870 nucleotides, therefore all nucleotide positions are in reference to this allele. *NAT1*3* differs by the substitution of a cytosine for an adenine at nucleotide position 1095, whilst retaining a thymine at position 1088. *NAT1*10* has a cytosine substituted for an adenine, and an adenine replaces a thymine at nucleotide positions 1095 and 1088, respectively. *NAT1*11* lacks 9 nucleotides in a tri-nucleotide repeat sequence upstream of nucleotide position 1088. *NAT1*11* also has three other point substitutions: a guanine substituted for an adenine resulting in a conserved amino acid change at nucleotide position 459, a

* Corresponding author: E. Sim, Tel. (01865) 271596; FAX (01865) 271853; E-mail; esim@worf.molbiol.ox.ac.uk

† *Abbreviations:* AS-PCR: allele specific-PCR; bp: base pair(s); DNA: deoxyribonucleic acid; NAT: arylamine *N*-acetyltransferase; PABA: *p*-aminobenzoic acid; PCR: polymerase chain reaction; RFLP: restriction fragment length polymorphism.

Received 28 April 1997; accepted 22 August 1997.

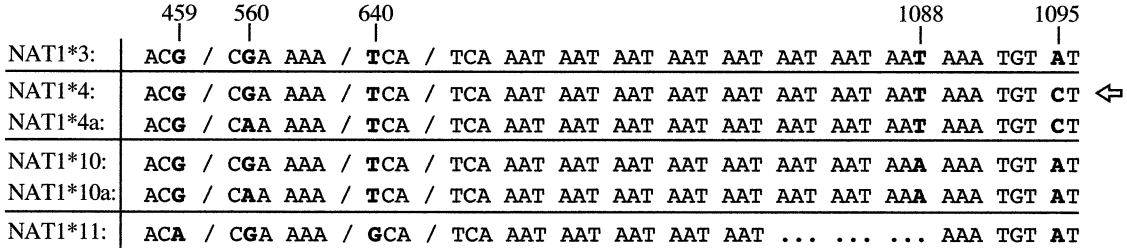

FIG. 1. Summary of mutations found in NAT1 alleles. Numbering is based on the first nucleotide of the initiation codon being represented as 1. The wild type allele is *NAT1*4* as indicated by arrows, nucleotides which are different to the wild type are shown in bold face.

thymine for a guanine leading to the replacement of a serine residue for an alanine at position 640, and a cytosine for an adenine which occurs outside of the coding region at position 1095.

This report describes two new allelic variants which contain nucleotide substitutions within the coding region (Fig. 1). These were identified by sequencing cDNA PCR amplified from gDNA representing NAT1 of an individual with low NAT1 activity assayed using *p*-aminobenzoic acid (PABA) as a substrate. We further describe an effective protocol for NAT1 genotyping which detects all described NAT1 alleles including the new allelic variants.

MATERIALS AND METHODS

Genomic DNA

Genomic DNA (gDNA) was extracted from whole blood [20] taken in EDTA from 16 healthy controls (18–32 years) and 16 patients with cystic fibrosis (17–37 years). All were non-related Caucasian volunteers [1].

Genotyping of Arylamine N*-Acetyltransferase Type 1*

Genotyping was determined by polymerase chain reaction-restriction fragment length polymorphisms (PCR-RFLP) and allele specific (AS)-PCR. PCR amplification was performed using 25 pmol of the appropriate primers designed to specifically anneal to the NAT1 isoform under investigation (Table 1). Amplification of cDNA was from 1 μg of genomic DNA or 50 ng of plasmid DNA per reaction using: 2.5 U of Taq polymerase (Promega), 2.0–2.5 mM $MgCl_2$, 0.2 mM of each dNTP and 1 × KCl reaction buffer (Promega). The reaction was made to a final volume of 100 μl with ultra pure water (Sigma) and overlaid with 50 μl of mineral oil (Sigma). The T_m of primers were calculated according to 67.5 + 34 (%GC expressed as a decimal)–(395/number of nucleotides). Annealing temperatures of primer pairs are shown in Table 1.

All PCR cycles were: a single step of 95° for 3 min ("hot start") followed by the addition of 2.5 U of Taq Polymerase, a denaturation step (95° for 30 sec), primer annealing (Table 1) and an elongation step at 72° based on 1 min per kilobase. The reaction consisted of 30–35 cycles and all reactions had a final elongation at 72° of 5 min.

Products from AS-PCR amplification and PCR-RFLP were visualised by gel electrophoresis with either 0.5 × TBE agarose (Boehringer Mannheim) and/or Metaphor (FMC Bioproducts). Primers N769 and N1113 (Table 1), and associated PCR-RFLP analysis, have been reported elsewhere [21].

Sequencing of NAT1 Allelic Variants

To confirm the validity of the genotyping strategy, six samples of gDNA from individuals with predetermined genotypes were PCR amplified, cloned and sequenced as follows. PCR amplification from gDNA was performed

TABLE 1. Primers for NAT1 genotyping

Primer	Primer sequence (5′–3′)	Sense/ Antisense	Annealing temperature
N-376:	(−376)/TAT TGC ATG ATT CTC CTG CCT A/(−355)	S	58°
N1177:	(1177)/GGA ATT CAA CAA TAA ACC AAC AT/(1155)	AS	
N769:	(769)/ACT CTG AGT GAG GTA GAA ATA/(789)	S	50°
N1113:	(1113)/ACA GGC CAT CTT TAG AA/(1096)	AS	(Bell *et al.*, [18])
N1110a:	(1110)/GGC CAT CTT TAA AAT ACA TTT A/(1089)	AS	54°
N1110b:	(1110)/GGC CAT CTT TAA AAT ACA TTT T/(1089)	AS	(used with N769)
N539:	(539)/TCC TAG AAG ACA GCA ACG ACC/(559)	S	52°
N714:	(714)/GTG AAG CCC ACC AAA CAG/(697)	AS	

Primers N769 and N1113 have been previously described (formerly named N1208 and N1536, Bell *et al.* [21]). Nucleotide numbering is with reference to the NAT1 open reading frame, the adenosine of the initiation codon representing 1. Nucleotide site of primer annealing is indicated in brackets. Primer N539 contains a cytosine (C) and a guanine (G) substituted for an adenosine (A) and thymine (T), at nucleotide positions 555 and 556 respectively, to create a Bsa OI restriction site (CGAC ⇓ CG) in alleles without the novel mutation. All nucleotide substitutions within primer sequences are depicted by an underscore.

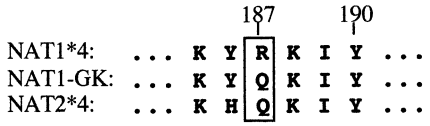

FIG. 2. Protein sequence comparison between human NAT1 and NAT2 wild types to that predicted for individual GK. Both alleles of GK are predicted to encode for the same protein resulting in an arginine (R) for a glutamine (Q) substitution.

using high fidelity proof reading pfu DNA polymerase (error rate of 1.3×10^{-6}; Stratagene) and the primers N376 and N1177 (Table 1). Amplified cDNA was 3′ tagged with dATP and cloned into the vector pGEMT (Promega). The entire cDNA product, 1552 bp from alleles *NAT1*3*, *NAT1*4* and *NAT1*10*, and 1543 bp from *NAT1*11* alleles, of these ten allelic variants were analysed on an ABI 377 automated sequencer by primer walking (Department of Biochemistry DNA sequencing service, University of Oxford). At least 2 clones were sequenced for well characterised genotypes such as *NAT1*4* and up to 8 clones were sequenced for new or rarer alleles.

Arylamine N*-Acetyltransferase Type 1 Activity*

Arylamine *N*-acetyltransferase type 1 activity was predetermined from frozen erythrocyte lysates using 1 mM PABA and 440 μm acetyl Coenzyme A as substrates, as previously described [1]. NAT1 activities derived from the genotypes *NAT1*4* homozygous and *NAT1*4/1*10* were then compared.

RESULTS

Identification of New NAT1 Alleles

Automated sequencing confirmed that genotypes obtained for previously described alleles [19] were as calculated except one. Cloned PCR products gave expected sequences at all nucleotide positions except GK. Individual GK, whose gDNA was genotyped as *NAT1*4/1*10*, contained an additional substitution within the coding region at nucleotide position 560 on both alleles when compared to the wild type *NAT1*4* (Fig. 1). Identical sequence was obtained from a further 4 clones representing each allele. A substitution of an adenine for the wild type guanine results in an amino acid change from an arginine residue to glutamine at amino acid 187 (Fig. 2). These new allelic variants were detected by the construction of a primer (N539) which possessed a partial Bsa OI restriction site. PCR-RFLP was performed using primers N539 and N714 (Table 1) followed by restriction digestion using Bsa OI. A series of gDNA samples from 32 individuals were amplified and digested. Results from 4 of these individuals are shown in Fig. 3. Non-mutated alleles (individuals CS, BD and RB) gave fragments after restriction digestion of 155 bp and 20 bp, mutated alleles (individual GK) interrupted the restriction site to yield a single product of 175 bp (Fig. 3). The other 28 samples gave the same genotype as did individuals

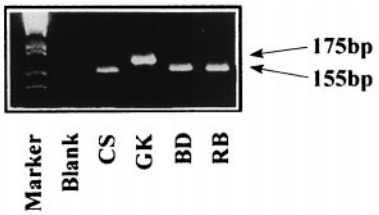

FIG. 3. PCR-RFLP analysis to identify the novel alleles *NAT1*4a* and *NAT1*10a*. Annealing of primers N539 and N714 was at 52°, elongation for 30 sec at 72° and 35 cycles were performed with 2.0 mM $MgCl_2$. Products from gDNA of four individuals is illustrated. The blank contained no gDNA and products were aligned against a 1 kb marker and analysed on a 4% Metaphor gel. Using the method outlined in Fig. 4, individuals GK and RB were genotyped as *NAT1*4/1*10*, and individuals CS and BD as *NAT1*4/1*4*. The novel mutation does not possess a Bsa OI restriction site and hence yields a 175 bp product (GK). Alleles without this mutation yield products of 155 bp and 20 bp, (CS, BD and RB).

CS, BD and RB. Individual GK possessed two mutant alleles *NAT1*4a* and *NAT1*10a*, no other individuals possessed these novel allelic variants. *NAT1*4a* is defined as the wild type allele with a nucleotide substitution at 560 (G ⇨ A). The allelic variant *NAT1*10a* is defined as a point mutation at 560 (G ⇨ A) associated with nucleotide substitutions in the 3′ non coding region (T ⇨ A) and 1095 (C ⇨ A) when compared to the wild type allele *NAT1*4* (Fig. 1).

NAT1 Genotyping from gDNA

NAT1 genotypes were determined from gDNA extracted from whole blood. A step-wise investigation was performed to identify the previously described NAT1 genotypes (Fig. 4, Table 2). As detailed by Bell *et al.* [21], PCR-RFLP analysis was performed in which cDNA was amplified using primers N769 and N1113 (Table 1) to yield an expected product of 344 bp from alleles *NAT1*3*, *NAT1*10* and *NAT1*4*, and 335 bp from *NAT1*11* alleles. The N1113 primer contains a partial Mbo II restriction site which forms a complete site on amplification from the *NAT1*4* allele. The PCR products were restriction digested with Mbo II for at least 2 hr at 37° and visualised by electrophoresis (Fig. 4a). From Fig. 4a, it can be seen that samples 2 and 6 are homozygous for *NAT1*4*, and thus no further genotyping was required. However, samples 3, 4, 5 and 7 may be either *NAT1*3* or *NAT1*10* and hence required additional genotyping. A second PCR-RFLP was performed to confirm sample 1 as *NAT1*4/NAT1*11* (Table 2).

The remaining samples 1, 3, 4, 5 and 7 were then analysed by AS-PCR (Fig. 4b). This was performed using either primers N1110a or N1110b in combination with N769 (Table 1). Primers N1110a and N1110b were designed specifically to distinguish the point mutation present at position 1088 resulting in a thymine for adenine substitution (Fig. 1). PCR amplification using the primer set N1110b and N769 generates products specific to *NAT1*10* alleles only. However, PCR amplification using primers N1110a and N769 will generate products from *NAT1*3*,

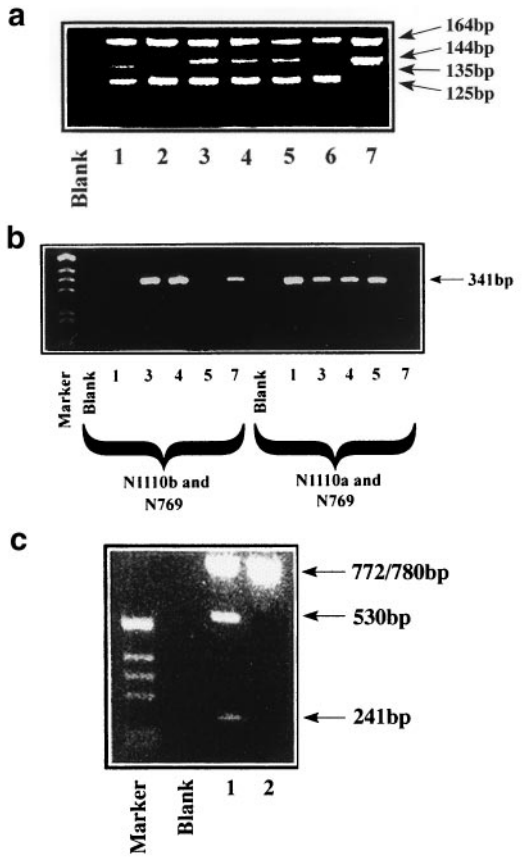

FIG. 4. Step-wise analysis of NAT1 genotypes. Fragment sizes were determined from a 1 kb ladder (Gibco BRL) and the blank contained no gDNA template. Samples 1–7 are gDNA from control individuals, sample 3 represents individual GK. All primers used are described in Table 1. (a) PCR-RFLP analysis was performed using primers N769 and N1113 followed by digestion with Mbo II, as previously described [21]. Products were analysed on a 4% Metaphor gel. Fragments of 164 bp and 124 bp were estimated for *NAT1*4*, 164 bp and 144 bp for *NAT1*3* and *NAT1*10*, and 164 bp and 135 bp for *NAT1*11* due to a 9 bp deletion (Fig. 1). (b) Differentiation of the *NAT1*3* and *NAT1*10* genotypes by AS-PCR. gDNA from samples 1, 3, 4, 5 and 7 were PCR amplified using primers N769 with either N1110a (will not anneal to *NAT1*10* alleles) or N1110b (will only anneal to *NAT1*10* alleles). Annealing was at 54°, elongation for 45 sec and 30 cycles were performed with 2.5 mM $MgCl_2$ to yield a 341 bp product analysed on a 2% Metaphor/1% agarose gel. (c) Confirmation of *NAT1*11* by PCR-RFLP analysis using primers N-376 and N1177, followed by digestion with A1w NI. Annealing was at 58°, elongation for 2 min and 35 cycles with 2.0 mM $MgCl_2$. *NAT1*11* alleles possess an additional A1w NI site to yield fragment sizes estimated as 772 bp, 530 bp and 241 bp compared to non-*NAT1*11* alleles which yield two fragments of 772 bp and 780 bp, as analysed on a 2% agarose gel. The calculated genotypes are: *NAT1*4/1*4* for samples 2 and 6, *NAT1*4/1*3* for sample 5, *NAT1*4/1*10* for samples 3 and 4, *NAT1*10/1*10* for sample 7, and *NAT1*4/1*11* for sample 1.

*NAT1*4* and *NAT1*11* alleles, but will not amplify from *NAT1*10* alleles. Thus, samples 3, 4 and 7 possessed a *NAT1*10* allele. Sample 7 was homozygous *NAT1*10* as no amplification occurred with N1110a and N769, and sample 5 must possess a *NAT1*3* allele as there was no amplification with N1110b and N769.

To confirm the presence of a *NAT1*11* allele in sample 1, initially implied from Fig. 4a, an additional PCR-RFLP analysis was performed in which gDNA was amplified using primers N-376 and N1177 to yield a predicted product of 1552 bp from alleles *NAT1*3*, *NAT1*4* and *NAT1*10*, and 1543 bp from *NAT1*11* alleles. A non-*NAT1*11* allele contains one A1w NI restriction site to reveal fragments on digestion of 772 bp and 780 bp. Whereas cDNA amplified from a *NAT1*11* allele contains an additional A1w NI site to yield fragment sizes of 772 bp, 530 bp and 241 bp (241 bp fragment appears faint in Fig. 4c). This confirmed the presence of a *NAT1*11* allele in sample 1, but not sample 2.

Comparison of NAT1 Genotype and Phenotype

Of the thirty-two individuals genotyped, eighteen were *NAT1*4* homozygous, ten were *NAT1*4/1*10*, and one of each of the following genotypes were identified: *NAT1*4a/1*10a*, *NAT1*4/1*11*, *NAT1*4/1*3* and *NAT1*10* homozygous. NAT1 activities from erythrocyte preparations were available for all genotyped samples. No significant differences in genotype distribution or erythrocyte NAT1 activities were observed between the two sample populations; normal individuals and patients with cystic fibrosis. Data from the two populations were pooled and NAT1 activities from individuals genotyped *NAT1*4* homozygous and *NAT1*4/NAT1*10* were analysed (Fig. 5). As too few individuals possessed other genotypes, these were ignored, though one individual: genotyped *NAT1*4/1*11*, had a relatively low NAT1 activity of 0.62. Mean activities, represented as a proportion of a control pooled population, for *NAT1*4/1*4* and *NAT1*4/1*10* were 1.131 (1.077 including *NAT1*4a/1*10a*) and 1.032, respectively. No significant difference in activities between the two groups *NAT1*4/1*4* and *NAT1*4/1*10* was observed ($P = 0.66$ with *NAT1*4a/1*10a*, $P = 0.31$ without *NAT1*4a/1*10a*; Student's *t*-test,) though a wide spread of activity, 0.53 (GK) to 1.59, was noted. Apart from three individuals, the NAT1 activities appeared to be clustered in a central group. Individual GK, *NAT1*4a/1*10a*, accounts for a low activity (indicated by a cross, Fig. 5) and two other individuals, genotyped *NAT1*4* homozygous, have higher activities (indicated by triangles, Fig. 5) when compared to the population as a whole. The cluster of *NAT1*4/1*10* has a higher average activity (1.13) than the central cluster of *NAT1*4/1*4* (0.89).

DISCUSSION

This report describes the identification of two NAT1 allelic variants (Fig. 1), and the development of an effective protocol for genotyping the known NAT1 alleles from gDNA prepared from whole blood. The efficacy of this protocol was confirmed by sequence analysis of the corresponding NAT1 alleles. The described genotyping protocol

TABLE 2. Step-wise investigation of described NAT1 genotypes

Step	Primers used	Products obtained	Next step
1	N539 and N714	155 bp = *NAT1*3*, *1*4*, *1*10* or *1*11*	Step 2
		175 bp = *NAT1*4a* or *1*10a*	Step 2
2	N769 and N1113	Mbo II digestion; 164 bp + 124 bp = *NAT1*4*	End
		164 bp + 144 bp = *NAT1*3* or *1*10*	Step 3
		164 bp + 135 bp = *NAT1*11*	Step 4
3	N769 and N1110a	341 bp = *NAT1*3*	End
	N769 and N1110b	341 bp = *NAT1*10*	End
4	N-376 and N1177	AlwNI digestion; 772 bp + 780 bp = *NAT1*3*, *1*4* and *1*10*	End
		773 bp, 530 bp + 241 bp = *NAT1*11*	End

Primer sequences are depicted in Table 1. Results of PCR amplifications and PCR-RFLP can be seen in Figs. 3 and 4.

has been applied with success to cells collected from the bladder as barbotage fluid or from mouth washings. According to previous NAT2 nomenclature [19], of which many alleles have been described, variants of similar alleles were named with a suffix A, B, C. Therefore, the novel variants described are termed *NAT1*4a* and *NAT1*10a* due to their similarity to the respective alleles. Since the initial submission of this report, Grant *et al.* [22] have also described the allele *NAT1*10a* and termed it *NAT1*14*. This further supports the identification of one of the two novel alleles.

The substitution occurring within both alleles (*NAT1*4a* and *NAT1*10a*) for individual GK, at nucleotide position 560, results in an arginine, a basic residue, being replaced by a glutamine, an amide (Fig. 2). NAT1 activities using PABA as a NAT1 substrate revealed erythrocytes isolated from GK to have a low acetylation phenotype compared to other controls (Fig. 5). The amino acid substitution found in GK has been suggested to be associated with arylamine specificity [23]. Moreover, this alteration matches that found with the NAT2 isoenzyme within the corresponding protein sequence (Fig. 2). The NAT2 isoenzyme does not acetylate PABA as a substrate and it has been suggested that the arginine/glutamine substitution contributes to the difference in the specificity of the two NAT isoforms [23]. The substitution in both alleles of GK is therefore likely to contribute to the low PABA activity. Preliminary *in vitro* studies with *NAT1*4a* expressed in *E. coli* confirm this suggestion as do activity measurements in cytosol from intestinal mucosa of an individual genotyped as carrying the *NAT1*10a* (*NAT1*14*) allele [24]. Further *in vitro* and *in vivo* studies are needed to determine the significance of this amino acid substitution. Only the *NAT1*11* allele has so far been reported to contain substitutions within the coding region at nucleotide positions 459 and 640. The substitution at 459 does not affect the encoded amino acid and the latter encodes for an amino acid change of a serine for an alanine.

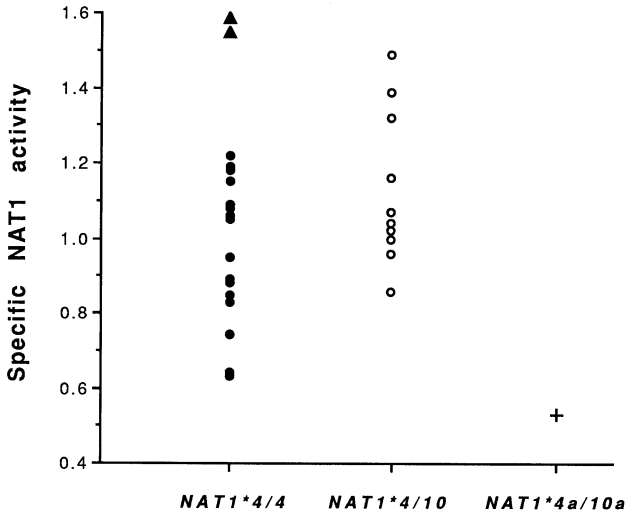

FIG. 5. Comparison of NAT1 genotypes to erythrocyte *N*-acetyltransferase activity. Activity is expressed relative to pooled erythrocyte samples measured at the same time, the average represented as 1.0 (Risch *et al.* [1]). Individuals genotyped as *NAT1*4/1*10* (open circles), *NAT1*4* homozygous (closed circles or triangles) and individual GK (cross), genotyped as *NAT1*4a/1*10a*, are shown.

It has been previously suggested that individuals heterozygous for *NAT1*10* had on average a two fold higher specific activity in bladder and colon homogenates using PABA as a substrate on comparison to *NAT1*4* homozygous individuals [10, 21]. Complementary to this, a significant association with the development of colorectal cancer and the possession of at least one *NAT1*10* allele was also reported [25]. However, on studying a larger case-control group, others observed no association between the possession of a *NAT1*10* allele and colorectal adenomas [26]. Contradictions in findings such as these may be explained by the presence of alleles *NAT1*4a* and/or *NAT1*10a*, or other new allelic variants. When individuals at the extremes of the acetylation range are removed, the results presented in Fig. 5 demonstrate a trend towards a higher NAT1 activity with individuals genotyped as *NAT1*4/1*10* compared to *NAT1*4* homozygous individuals, though the difference is not two fold as reported by others [10, 21]. This discrepancy may be because activities in this report derive from erythrocytes whereas cytosol from other tissues were used in previous studies.

The amino acid substitution reported here in the alleles of individual GK appears to result in low NAT1 activity. This novel allelic variant appears to be rare: only individual GK possessed this nucleotide substitution of the thirty-two individuals genotyped. Though activity is low in individual GK, it was detectable and to date no report exists on

individuals which lack NAT1 activity. It has been suggested that NAT1, which has a widespread tissue distribution, may play a role in folate catabolism [27]. Folate is catabolised to *p*-aminobenzoyl glutamate (*p*ABAGlu; an endogenous substrate for NAT1) which is in turn *N*-acetylated to *N*-acetyl *p*ABAGlu. Levels of erythrocyte folate are high in individual GK [28] which may provide further evidence for NAT1 having a role in folate catabolism.

Further investigations are required to determine whether phenotypic variations, if due to NAT1 alone, are the result of nucleotide differences outside of the coding region or amino acid substitutions within the open reading frame. Investigations to date have concentrated on nucleotide substitutions within the polyadenylation signal, but it is likely that other transcriptional regulatory elements require examination to determine whether mRNA stability, as previously implied by others [25], contributes to inter-individual variation in NAT1 activity.

We are grateful to all colleagues and patients who gave blood, and to the Wellcome Trust for generous financial support. We thank Valerie Smelt for helpful discussion. We are also grateful for the excellent technical assistance of Mr. Jason Pope and Mrs. Aileen Watson.

References

1. Risch A, Smelt V, Lane D, Stanely L, van der Slot W, Ward A and Sim E, Arylamine *N*-acetyltransferase in erythrocytes of cystic fibrosis patients. *Pharmacol Toxicol* **78:** 235–240, 1996.
2. Giovannucci E, Colditz GA, Stampfer MJ, Hunter D, Rosner BA, Willett WC and Speizer FE, A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in U.S. women. *J Natl Cancer Inst* **86:** 192–199, 1994.
3. Potter JD, Slattery ML, Bostick RM and Gapstur SM, Colon cancer: a review of the epidemiology. *Epidemiol Rev* **15:** 499–545, 1993.
4. Smith G, Stanley LA, Sim E, Strange RC and Wolf CR, Metabolic polymorphisms and cancer susceptibility. *Cancer Surv* **25:** 27–65, 1995.
5. The Office of Population Censuses and Surveys (OPCS). Cancer statistics registration—Registrations of cancer diagnosed in 1988, England and Wales. London: HMSO, 1994.
6. Talaska G, Schamer M, Skipper P, Tannenbaum S, Caporaso N, Kadlubar F, Bartsch H and Vineis P, Carcinogen-DNA adducts in exfoliated urothelial cells: techniques for noninvasive human monitoring. *Environ Health Perspect* **99:** 289–291, 1993.
7. Wallace DMA, Occupational urothelial cancer. *Br J Urol* **61:** 175–182, 1988.
8. Case RA, Hosker ME, McDonald DB and Pearson JT, Tumours of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. *Brit J Ind Med* **11:** 75–104, 1954.
9. Ambrosone CB, Freudenhein JL, Graham S, Marshall JR, Vena JE, Brasure JR, Michalek AM, Laughlin R, Nemeto T, Gillenwater KA, Harrington AM and Sheilds PG, Cigarette smoking, *N*-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. *JAMA* **276:** 1494–1501, 1996.
10. Badawi AF, Hirvonen A, Bell DA, Lang NP and Kadlubar FF, Role of aromatic amine acetyltransferases, NAT1 and NAT2, in carcinogen-DNA adduct formation in the human urinary bladder. *Cancer Res* **55:** 5230–5237, 1995.
11. Andres HH, Klem AJ, Schopfer LM, Harrison JK and Weber WW, On the active site of liver acetyl-CoA arylamine N-acetyltransferase from rapid acetylator rabbits (111/J). *J Biol Chem* **263:** 7521–7527, 1988.
12. Andres HH, Kolb HJ, Schreiber RJ and Weiss L, Characterisation of the active site, substrate specificity and kinetic properties of acetyl-CoA: arylamine *N*-acetyltransferase from pigeon liver. *Bioch Biophys Acta* **746:** 193–201, 1983.
13. Blum M, Grant DM, McBride W, Heim M and Meyer UA, Human arylamine *N*-acetyltransferase genes: isolation, chromosomal localization, and functional expression. *DNA Cell Biol* **9:** 193–203, 1990.
14. Matas N, Thygesen P, Stacey M, Risch A and Sim E, Mapping AAC1, AAC2 and AACP, the genes for arylamine *N*-acetyltransferases, carcinogen metabolising enzymes of human chromosome 8p22, a region frequently deleted in tumours. *Cytogen Cell Genet* **77:** 290–295, 1997.
15. Hearse DJ and Weber WW, Multiple *N*-acetyltransferases and drug metabolism. *Biochem J* **132:** 519–526, 1973.
16. Pacifici GM, Bencini C and Rane A, Acetyltransferase in humans: development and tissue distribution. *Pharmacology* **32:** 283–291, 1986.
17. Risch A, Wallace DM, Bathers S and Sim E, Slow *N*-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer. *Hum Mol Genet* **4:** 231–236, 1995.
18. Grant DM, Morike K, Eichelbaum M and Meyer UA, Acetylation pharmacogenetics. The slow acetylator phenotype is caused by decreased or absent arylamine *N*-acetyltransferase in human liver. *J Clin Invest* **85:** 968–972, 1990.
19. Vatsis KP, Weber WW, Bell DA, Dupret JM, Evans DA, Grant DM, Hein DW, Lin HJ, Meyer UA, Relling MV, Sim E, Suzuki T and Yamazoe Y, Nomenclature for *N*-acetyltransferases. *Pharmacogenetics* **5:** 1–17, 1995.
20. DiLella AG and Woo SL, Cloning large segments of genomic DNA using cosmid vectors. *Methods Enzymol* **152:** 199–212, 1987.
21. Bell DA, Stephens EA, Castranio T, Umbach DM, Watson M, Deakin M, Elder J, Hendrickse C, Duncan H and Strange RC, Polyadenylation polymorphism in the acetyltransferase 1 gene (NAT1) increases risk of colorectal cancer. *Cancer Res* **55:** 3537–3542, 1995.
22. Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A, Yu VL and Grewal R, Human acetyltransferase polymorphisms. *Mut Res* **376:** 61–70, 1997.
23. Sinclair J and Sim E, A fragment consisting of the first 204 amino-terminal amino acids of human arylamine *N*-acetyl transferase one (NAT1) and the first transacetylation step of catalysis. *Biochem Pharmacol* **53:** 11–16, 1997.
24. Hickman D, Pope J, Patil SD, Fakis G, Stanely LA, Payton M, Unadkat JP and Sim E, Expression of arylamine *N*-acetyltransferase in human intestine. *Gut*. (In press), 1997.
25. Bell DA, Badawi AF, Lang NP, Ilett KF, Kadlubar FF and Hirvonen A, Polymorphism in the *N*-acetyltransferase 1 (NAT1) polyadenylation signal: association of *NAT1*10* allele with higher *N*-acetylation activity in bladder and colon tissue. *Cancer Res* **55:** 5226–5229, 1995.
26. Probst-Hensch NM, Haile RW, Li DS, Sakamoto GT, Louie AD, Lin BK, Frankl HD, Lee ER and Lin HJ, Lack of association between the polyadenylation polymorphism in the *NAT1* (acetyltransferase 1) gene and colorectal adenomas. *Carcinogenesis* **17:** 2125–2129, 1996.
27. Minchin RF, Acetylation of para-aminobenzoylglutamate, a folate catabolite, by recombinant human NAT and U937 cells. *Biochem J* **307:** 1–3, 1995.
28. Ward A, Hickman D, Gordon JW and Sim E, Arylamine *N*-acetyltransferase in human red blood cells. *Biochem Pharmacol* **44:** 1099–1104, 1992.