CANCER

Polymorphisms of the human *UDP-glucuronosyltransferase (UGT) 1A7* gene in colorectal cancer

C P Strassburg, A Vogel, S Kneip, R H Tukey, M P Manns

....

See end of article for authors' affiliations
....

Correspondence to:
Dr C P Strassburg,
Department of
Gastroenterology and
Hepatology, Hannover
Medical School,
Carl-Neuberg-Str 1, 30625
Hannover, Germany;
strassburg.christian@
mh-hannover.de

Accepted for publication
7 August 2001
....

Background: Genetic polymorphisms in the human *UDP-glucuronosyltransferase-1A7* (*UGT1A7*) gene are detected and significantly correlated with sporadic colorectal carcinoma. UGT1A7, which has recently been demonstrated to glucuronidate environmental carcinogens, is now implicated as a cancer risk gene. A silent mutation at codon 11 and missense mutations at codons 129, 131, and 208 lead to the description of three polymorphic alleles designated UGT1A7*2, UGT1A7*3, and UGT1A7*4.

Methods: UGT1A7 polymorphisms were analysed by polymerase chain reaction amplification and sequencing, as well as temperature gradient gel electrophoresis in 210 healthy blood donors and 78 subjects with colorectal cancer.

Results: Homozygous wild-type UGT1A7 alleles were present in 20% of normal controls but were only detected in 9% of patients with colorectal carcinoma (odds ratio (OR) 0.39 (95% confidence interval (CI) 0.17–0.92); p=0.03). Analysis of individual polymorphic alleles identified a highly significant association between the presence of UGT1A7*3 alleles and colorectal cancer (OR 2.75 (95% CI 1.6 – 4.71); p<0.001). Recombinant expression of UGT1A7 polymorphic cDNA in eukaryotic cell culture showed reduced carcinogen glucuronidation activity in comparison with wild-type *UGT1A7*. *UGT1A7* may therefore represent a modifier gene in colorectal carcinogenesis.

Conclusion: We have identified a potential novel risk factor in sporadic colorectal cancer which may contribute to the identification of risk groups and to the elucidation of factors involved in colon carcinogenesis.

The human UDP-glucuronosyltransferase (UGT) UGT1A7[1–3] belongs to a family of nine proteins encoded at the human *UGT1A* gene locus on chromosome 2q37, which is highly conserved in vertebrate organisms.[4] UGTs catalyse the conversion of hydrophobic substrates to inactive hydrophobic β-D-glucopyranosiduronic acids (glucuronides) which can undergo facilitated renal or biliary elimination. Target substrates for glucuronidation consist of compounds from many divergent chemical classes, including dietary byproducts, endogenous metabolites such as steroids and bilirubin, many therapeutic drugs, and many environmental mutagens, such as polycyclic and heterocyclic aromatic hydrocarbons as well as heterocyclic amines.[4] These enzymes play a critical role in cellular defence and are considered to be a major class of proteins involved in cellular detoxification.

In humans, tissue specific expression of the UGT1A locus gene products has been demonstrated.[3 5] UGT1A7, UGT1A8, and UGT1A10 are not expressed in hepatic tissues and have been shown to be differentially regulated in gastrointestinal organs such as the oesophagus,[6] stomach,[3] small intestine,[2] and colon.[7] In addition, regulation and function of the *UGT1A* genes has been linked to gastrointestinal carcinogenesis.[2 4 6 8] Analysis of human cancers has shown differential regulation of UGT expression in tumour tissues with an impact on epithelial glucuronidation activity.[2 8] In gastric mucosa, *UGT1A* regulation leading to interindividual variations in UGT activity has been implicated as a cancer risk factor[2] while in oesophageal cancer, downregulation of UGT1A7 transcripts is correlated with a decrease in overall oesophageal benzo(α)pyrene metabolite glucuronidation.[6] This is significant as the cloning and characterisation of UGT1A7 has demonstrated its capacity to glucuronidate polycyclic aromatic hydrocarbons as well as dietary derived heterocyclic amines,[3 6] both of which are known carcinogens. These findings prompted us to predict that alterations in *UGT1A7* gene expression or in the functional properties of UGT1A7 may be an indicator of potential cancer risk. Hence we examined *UGT1A7* as a target gene for the evaluation of polymorphisms and their relation to gastrointestinal cancer.

Colorectal cancer (CRC) accounts for 9.7% of all newly diagnosed carcinomas.[9] Pedigree and genetic analyses have identified a heritable aetiology in 15–20% of colorectal cancers. This has been linked to mismatch repair genes which have thus been indicated as colon cancer susceptibility markers.[10–16] For a proportion of the remaining sporadic cancers, carcinogenesis based on mutagen mediated genotoxicity capable of initiating neoplastic transformation represents a potential mechanisms.[17] The combined risk of carcinogenesis may therefore be determined by the outcome of an interaction between genetically defined biochemical defence mechanisms of the cell, such as those provided by the well characterised UGT super gene family, and the environmentally defined exposure to carcinogens.[18 19]

In this study we show the significant association of CRC with polymorphisms of the UGT1A7 carcinogen metabolising *UGT* gene.

MATERIALS AND METHODS

Patients

Patient samples were collected from consecutive patients attending the Department of Gastroenterology and Hepatology, and Department of Transfusion Medicine, Hannover Medical School in 1999 and 2000. All patients were of German Caucasoid ancestry. Written consent was obtained from all

....

Abbreviations: UGT, UDP-glucuronosyltransferase; CRC, colorectal cancer; TGGE, temperature gradient gel electrophoresis; PhIP, amino-1-methyl-6-phenylimidazo-(4.5-β)-pyridine.

patients prior to blood sampling and the ethics committee of Hannover Medical School approved the study protocol. The following patient sample groups were studied. As controls, healthy blood donors were studied (n=210, females 92/males 118; mean (SD) age 48 (21) years (range 19–85)). Colorectal cancer patients (n=78, females 28/males 50; mean (SD) age 63.3 (11) years (range 38–85)) had sporadic colorectal cancer identified by exclusion of the Amsterdam II criteria.[20] The diagnosis was based on histological evaluation of endoscopic or surgical specimens in each patient.

Genomic DNA

Genomic DNA was prepared from full blood samples using the QiaAmp isolation system according to the recommendations of the manufacturer (Qiagen, Hilden, Germany). Concentrations were determined by spectrophotometry at 260 and 280 nm and samples were stored in 10 mM Tris/EDTA buffer (pH 8.0) at 4°C until analysis.

Analysis of UGT1A7 exon 1 sequences

The human *UGT1A* gene locus consists of divergent individually regulated exon 1 sequences at the 5′ end followed by one copy of exons 2–5 at the 3′ end. Each individual UGT1A transcript therefore comprises unique exon 1 sequences combined with exon 2–5 sequences present on all UGT1A transcripts. Therefore, in order to study a specific *UGT1A* gene, examination of exon 1 sequence is required. UGT1A7 exon 1 sequences were amplified by polymerase chain reaction. The forward primer was located from bp −61 to −38 upstream of the ATG start codon of UGT1A7 (GenBank accession No U39570) (5′-gcggctcgagccacttactatattataggagct-3′); the reverse primer was located between bp 855 and 829 (GenBank accession No U89507) of the UGT1A7 exon 1 sequence (5′-gcggatatccataggcactggctttccctgatgaca-3′). Location of the upstream primer outside the open reading frame was required to obtain specificity of the amplification reaction because exon 1 sequences of UGT1A7–10 share 93% homology. A 916 bp product was amplified in a volume of 100 µl containing 10 mM KCl, 20 mM Tris HCl (pH 8.8), 10 mM ammonium sulphate, 2 mM magnesium sulphate, 1% Triton X-100, 0.2 mM each dNTP, 20 ng of genomic DNA, 2 µM of primers, and 5 units of VENT (exo-) DNA polymerase (NEB, Beverly, Massachusetts, USA). After a hot start at 94°C for three minutes, 30 cycles of 94°C for 30 seconds, 57°C for 30 seconds, and 72°C for 90 seconds were run on a Perkin Elmer GeneAmp PCR 2400 system. The products were visualised by 2% agarose gel electrophoresis and purified using QuiaQuick Columns according to the manufacturer's recommendations (Qiagen). The sequences of the polymerase chain reaction products were determined by automated fluorescence sequencing (MWG-Biotech, Ebersbach, Germany). Sequence data were analysed using the PC Gene software package (Oxford Molecular, Campbell, California, USA).

Statistical analysis

Statistical analysis (two tailed Fisher's exact test, odds ratios (ORs), confidence intervals (CI)) was carried out using the EpiCalc 2000 software package, available on the worldwide web (www.myatt.demon.co.uk/epicalc.htm). Logistical regression analyses were calculated using the SPSS software package (SPSS Inc. Chicago, Illinois, USA) for all identified genotypes, sex, and the combined UGT1A7*3 allelic genotypes.

Temperature gradient gel electrophoresis

In addition to obtaining sequencing data, polymerase chain reaction products were analysed by temperature gradient gel electrophoresis (TGGE) to detect DNA heteroduplex formation. Two separate experiments were optimised to detect point mutations at codons 129/131 as well as at codon 208 using the

POLAND software package available on the worldwide web (www.biophys.uni-duesseldorf.de/POLAND/poland.html). A DNA fragment containing codons 129/131 was amplified with the sense primer 5′-caatggtatttttgactt-3′ and the antisense primer 5′-ggcggcggcggggcgcggcgggtaaatgttcctcc-3′, which was extended by a "gc clamp". Amplification was carried out for 30 cycles at 94°C (20 seconds), 58°C (30 seconds), and 72°C (30 seconds) using Taq DNA polymerase (Promega, Madison, Wisconsin, USA) under conditions recommended by the supplier. Gel electrophoresis was performed in a 10% polyacrylamide gel using the TGGE system (Biometra-Whatman, Göttingen, Germany) and a gradient of 25–30°C. Analysis of codon 208 mutations was performed by polymerase chain reaction amplification (sense primer: 5′-cccgccgcgccccgccgcccgccagtgccctgctcc-3′ which was extended by a "gc clamp", antisense primer: 5′-cagaggctatttctaaga-3′) for 30 cycles at 94°C (20 seconds), 61°C (30 seconds), and 72°C (30 seconds) using Taq DNA polymerase (Promega) under conditions recommended by the supplier. TGGE gel electrophoresis was carried out over a gradient of 39–46°C.

Site directed mutagenesis

Complementary DNA containing the polymorphic UGT1A7*4 (GenBank accession No AF296126), UGT1A7*2 (GenBank accession No AF296127), and UGT1A7*3 (GenBank accession No AF296128) alleles were generated by polymerase chain reaction based site directed mutagenesis. The recently cloned and published wild-type cDNA sequence (GenBank accession No U89507) was used as template DNA.[1] The mutagenesis strategy employed three primer pairs. The first pair (sense primer 1: 5′-ctgggatcc (BamHI)agttctctg*atg* (*start*)gctcgtgc-3′, antisense primer 2: cacaagctt(HindIII)*ctca* (*stop*) atgggtcttggatttgtgg-3′) spans the entire open reading frame of UGT1A7, includes restriction endonuclease sites for directional cloning, and has previously been published.[1] The second primer pair anneals to the sense and antisense strand between base pairs 370 and 405 covering the missense base pair changes at codon 129 and at codon 131 (sense primer 3: 5′-gcaggagtttgtttaa**g**gac**aa**aaaattagtagaa-3′, antisense primer 4: 5′-ttctactaatttt**tt**gtc**c**ttaaacaaactcctgc-3′). The third primer pair anneals to the sense and antisense strand between base pairs 603 and 637 covering the missense base pair changes at codon 208 (sense primer 5: 5′-gactttcaaggagagagta**c**ggaaccacatcatgc-3′, antisense primer 6: 5′-gcatgatgtggttcc**g**tactctctccttgaaagtc-3′). The UGT1A7*4 polymorphism was generated by separate amplification of the fragments generated with primers 1 and 6 as well as 5 and 2. After annealing of the denatured fragments in the region covered by primers 5 and 6 and an extension reaction, the full length product of 1593 bp was amplified using primers 1 and 2. The UGT1A7*2 polymorphism was generated in a similar fashion using primers 1 and 4 as well as 3 and 2 for the initial amplification and primers 1 and 2 to amplify the 1593 bp product after the extension reaction. The UGT1A7*3 polymorphism which combines missense mutations at codons 129, 131, and 208 required two rounds of annealing and amplification. Firstly, three fragments were generated by polymerase chain reaction using primers 1 and 4, primers 3 and 6, and primers 5 and 2. Next, the denatured fragments between primers 3 and 6 as well as between primers 5 and 2 were annealed for the extension reaction and subsequently amplified with primers 3 and 2. This product was denatured and annealed with the fragment between primers 1 and 4. The final amplification using primers 1 and 2 resulted in a full length product of 1593 base pairs. The products were purified after amplification to eliminate contamination with primers and template. Polymerase chain reaction was performed under the conditions described for the UGT1A7 exon 1 amplification above. All sequences were confirmed by sequencing of both strands. The wild-type UGT1A7*1 as well as the UGT1A7*2, UGT1A7*3, and UGT1A7*4 cDNAs were inserted into the multicloning site of
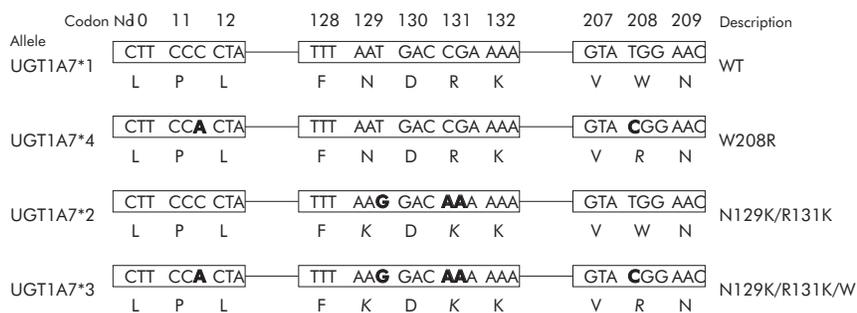
Figure 1 Graphic representation of the three polymorphic sites identified on the human *UGT1A7* exon 1 sequence. The mutation of CCC/A at codon 11 is silent. At codons 129, 131, and 208, mutations lead to amino acid substitutions. Sequence analysis identified the presence of three different polymorphisms designated UGT1A7*2, UGT1A7*3, and UGT1A7*4. Exchanges at codons 11 and 208 (UGT1A7*4) as well as the polymorphisms at codons 129 and 131 (UGT1A7*2) did not occur individually (GenBank Nos AF296226, AF292627, AF292628). Base pair exchanges are shown in bold type and amino acid exchanges in italics.

Figure 2 Compared with the region around codon 11 (left box), the identified polymorphisms at positions 129/131 and at position 208 fall within regions of highly conserved sequences (middle and right boxes). The N129K/R131K polymorphism represents the wild-type sequence of UGT1A9 at these positions. W208R represents the wild-type sequence of UGT1A8 and UGT1A9 at this position.

pCDNA3.1 (Invitrogen, San Diego, California, USA). The sequence of the generated polymorphic UGT1A7 cDNAs was determined by automated fluorescence sequencing. In addition, each product was analysed by in vitro transcription/translation (TNT quick Transcription/Translation Systems, Promega, Madison, Wisconsin, USA) to document generation of 55 kDa products for each clone prior to expression in HEK293 cells.

Expression and characterisation of UGT1A7*1, UGT1A7*2, UGT1A7*3, and UGT1A7*4

Human embryonic kidney (HEK293) cells were transfected with pCDNA3.1 plasmids containing the UGT1A7*1, UGT1A7*3, and UGT1A7*4 cDNAs as previously described.[3] Cells were harvested after 72 hours and lysed by freeze thawing. Expression of equal amounts of the different recombinant UGT proteins was controlled by β-galactosidase transfection and western blot using a rabbit anti-exon 2 antibody as reported previously.[3] The following substrates were resuspended in methanol at a concentration of 5 mM: 4-nitrophenol, 1-naphthol, 4-methylumbelliferone, 4-tert-butylphenol, 7-hydroxybenzo(α)pyrene, 8-hydroxybenzo(α)pyrene, 7/8-hydroxybenzo(α)pyrene, 9-hydroxybenzo(α)pyrene, and amino-1-methyl-6-phenylimidazo-(4.5-β)-pyridine (PhIP). Whole cell extracts (50 µg) expressing recombinant protein in reaction buffer (50 mM Tris-Cl (pH 7.6), 10 mM $MgCl_2$) were used for catalytic activity assays in a modification of the method of Bansal and Gessner,[21] as previously described.[5] The UDP-glucuronic acid concentration was 250 µM. The ^{14}C-labelled glucuronides were detected by thin layer liquid chromatography, visualised by autoradiography and quantitated using a Fujifilm BAS-1000 phosphoimager (Raytest GmbH, Straubenhardt, Germany) and TINA 2.0 software (Raytest GmbH, Straubenhardt, Germany) for calculation of specific activities that were expressed as pmol glucuronide formed/min/mg of recombinant

protein. Autoradiography hard copies were additionally analysed with a GS-710 calibrated imaging densitometer (BioRad Laboratories, Hercules, California, USA) using the Quantity One software package.

RESULTS AND DISCUSSION

The risk of cancer development is defined by genetic predisposition and environmental mechanisms such as exposure to cellular mutagens.[19 22] The cellular interaction with putative carcinogens is influenced by the activity of detoxification enzymes expressed in tissues with xenobiotic contact.[23 24] Therefore, the link between genetic polymorphisms of carcinogen metabolising enzymes and human cancer represents an attractive model combining genetic predisposition and environmental exposure.[4 17]

Detection of three polymorphic sites within exon 1 of the *UGT1A7* gene

In this study, UGT1A7 polymorphisms were analysed in 288 Northern German Caucasoid patients, which included 210 healthy controls and 78 patients with CRC. Confirming an earlier report analysing American patients,[25] three polymorphic sites with a total of five individual base pair changes leading to three amino acid changes were detected (figs 1, 2). The sequencing data were confirmed by a TGGE assay developed for the detection of the 129K/131K as well as the 208R missense mutations (fig 3). The combined differences are represented on three alleles defined as UGT1A7*2 (GenBank accession No (acn) AF296127), UGT1A7*3 (GenBank acn AF296128), and UGT1A7*4 (GenBank acn AF296126).

The UGT1A7*4 allele is characterised by a silent mutation within the endoplasmic reticulum signal peptide domain[1 4 26] which is present in all UGT1A proteins with the exception of
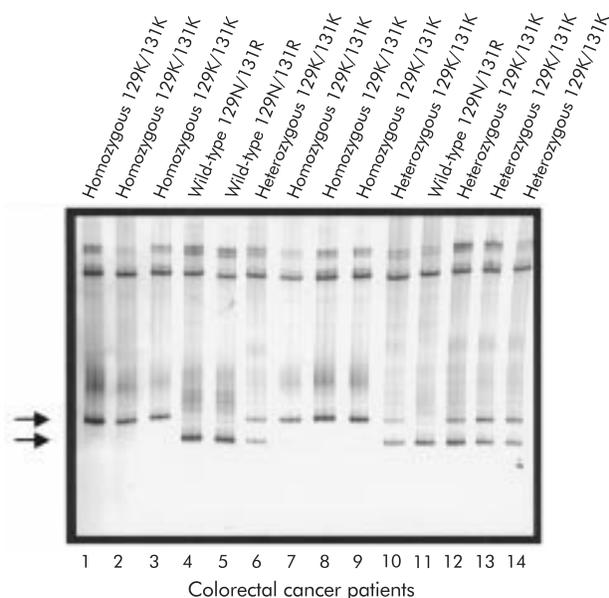
Figure 3 Detection of UGT1A7 129K/131K polymorphisms by temperature gradient gel electrophoresis. The silver stained 10% acrylamide gel separated over a temperature gradient of 25–30°C is shown. Heteroduplex formation leads to the distinction of two bands which allows for the discrimination of wild-type 129N/131R (lower band, lanes 4, 5, and 11), of the heterozygous 129K/131K (lower and upper bands, lanes 6, 10, and 12–14), and the homozygous mutant 129K/131K (UGT1A7*3/*3) (upper band only, lanes 1–3 and 7–9). Detection of the 208R polymorphism was performed similarly (not shown).

UGT1A10. The second change is represented by the missense mutation at codon 208. Codon 208 falls within a region of sequence conserved not only in the first exons of UGT1A7–10 but in all members of the UGT1A family (fig 2). The presence of an RV-N motif between amino acids 206 and 209 is common to all UGT1A proteins.

UGT1A7*2 combines two missense mutations in close proximity. As found for UGT1A7*4 this polymorphism also affects highly conserved sequence of the UGT protein. Following leucine-127, which is present in all UGT1A proteins, a LHN—L motif (amino acids 128–133) is conserved in UGT1A1, UGT1A3, UGT1A4, and UGT1A5, and a F-D-KLV motif (amino acids 128–134) is conserved in the extrahepatic

UGT1A7, UGT1A8, UGT1A9, and UGT1A10 (fig 2). UGT1A6 is shown to be more divergent at this site (LQDRDTL). The homologies at the polymorphic site are in agreement with the previously described overall homology within the UGT1A3–5 and the UGT1A7–10 clusters of UGT sequences.[3]

UGT1A7*3 is represented by the combination of all of the changes observed in UGT1A7*4 and UGT1A7*2. In our analysis, none of the 288 individuals demonstrated the single heterozygous or homozygous mutation at any of the individual codons displayed by the three alleles.

Significant association of UGT1A7 polymorphisms with colorectal carcinoma

Following the identification of polymorphic *UGT1A7* alleles, the occurrence of the three missense mutations in normal controls as well as in CRC was determined (table 1). We used a control group of 210 healthy blood donors in addition to a cancer group of 78 CRC patients. The number of patients studied resulted in a statistical power of 90% in the present study. In the control population, homozygosity for wild-type UGT1A7 was confirmed in 20%. Of the other UGT1A7 alleles, UGT1A7*2/*1 was found in 25%, UGT1A7*2/*2 in 12%, and UGT1A7*3/*3 in 10% of individuals. The occurrence of at least one allele of UGT1A7*3 was found to be present in 27% and the allelic frequency was 0.17 (table 1). In a previously reported series, 144 healthy blood donors from the Boston area were analysed and found to exhibit an allelic frequency of 0.361 for the UGT1A7*3 polymorphic allele[25] which contrasts with our findings in a homogeneous group of 210 German blood donors all of Caucasian descent from Northern Germany. Unfortunately, the study by Guillemette et al did not provide data on subjects' genetic backgrounds. It is therefore possible that differences in genetic/ethnic background and immigration patterns in the USA versus a more homogeneous population in Northern Germany contribute to the observed differences. In the present study, we compared a defined control group with patients with manifest CRC who were also of German descent from the same area, and therefore shared the genetic background.

In the CRC group, a striking difference in the occurrence of UGT1A7 polymorphisms was detected (table 1, fig 4). Only 9% of CRC patients exhibited the UGT1A7 homozygous wild-type (p=0.03, OR 0.39 (95% CI 0.17–0.92)). The high incidence of UGT1A7 polymorphisms detected among patients with sporadic CRC patients was remarkable and exceeded that of other individual genetic CRC markers identified to date.[27 28] In particular, the significance of this novel association differed

Table 1 Association of UGT1A7 polymorphisms with colorectal carcinoma

UGT1A7 alleles	Normal controls† (n=210) (%)	Colorectal cancer† (n=78) (%)	p Value	OR (95% CI)
UGT1A7*1/*1	42 (20)	7 (9)	0.03	0.39 (0.17–0.92)
UGT1A7*2/*1	52 (25)	13 (17)	0.15	
UGT1A7*2/*2	25 (12)	4 (5)	0.12	
UGT1A7*3/*1	20 (10)	15 (19)	0.04	2.26 (1.09–4.68)
UGT1A7*3/*2	19 (9)	15 (19)	0.02	2.39 (1.15–4.99)
UGT1A7*3/*3	17 (8)	8 (10)	0.63	
UGT1A7*4/*1	35 (17)	15 (19)	0.60	
UGT1A7*4/*3	0	1 (1)	NS	
UGT1A7*4/*4	0	0	NS	
Presence of UGT1A7*3	56 (27)	39 (50)	<0.001	2.75 (1.6–4.71)

The wild-type allele of UGT1A7 was significantly more prevalent in the tumour free control group (OR 0.39, 95% CI 0.17–0.92). The association of polymorphic UGT1A7 alleles with colorectal cancer was significant for the UGT1A7*3/*1 and the UGT1A7*3/*2 genotypes. When the data were analysed for the presence of UGT1A7*3 alleles a highly significant association with colorectal cancer (CRC) was found. This analysis identified the presence of the UGT1A7*3 allele as a risk factor for CRC. The frequency of the UGT1A7*3 allele was 0.17 in normal controls and 0.301 in CRC patients.
p values calculated by two tailed Fisher's exact test.
OR, odds ratio; CI, confidence interval; NS, not significant.
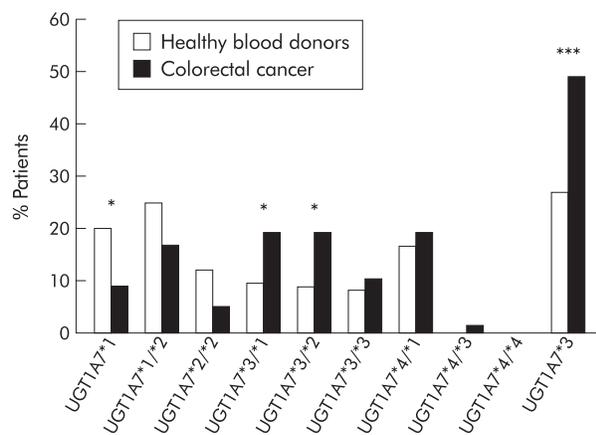†All patients were of Northern German Caucasoid descent.

Figure 4 Graphic representation of the association of polymorphic UGT1A7 alleles with sporadic colorectal cancer. A highly significant association was apparent for the wild-type alleles in normal control subjects without tumour and for the UGT1A7*3 allele with sporadic colorectal cancer. All numbers are in per cent (compare with table 1). All controls and colorectal cancer patients were of Northern German Caucasoid descent. Statistical analysis (two tailed Fisher's exact test): *p<0.05, ***p<0.001.

from that reported for other allelic associations of metabolising enzymes and cancer.[29][30]

When these data were further analysed for the presence of the different defined genotypes (table 1, fig 4), as characterised in fig 1, a statistically significant genotypic association with CRC was observed with UGT1A7*3/*1 as well as with UGT1A7*3/*2. This analysis identified the UGT1A7*3 allele as a risk factor in CRC. The allelic frequency of UGT1A7*3 in sporadic CRC was 0.30. Thus when the presence of UGT1A7*3 alleles was compared in normal control subjects and CRC patients, this difference was highly significant (p<0.001, OR 2.75 (95% CI 1.6–4.71)). Additionally, logistical regression analysis demonstrated that tumour occurrence was dependent on the presence of polymorphic UGT1A7 alleles (p=0.025), in particular UGT1A7*3 (p=0.001), but was independent of sex (p=0.17).

Modification of catalytic UDP-glucuronosyltransferase 1A7 activity

To our knowledge, these data provide the first evidence of a significant genetic link between the gene of a carcinogen metabolising enzyme and sporadic CRC. These findings raise significant aetiological implications for the disease. In sporadic CRC, environmental as well as genetic factors have been suggested, yet never directly demonstrated. An attractive hypothesis would suggest that alteration of xenobiotic or carcinogen glucuronidation in individuals expressing polymorphic alleles would affect exposure with these agents to the proximal as well as the downstream intestinal tract which could determine the rate of inflammation as well as cytotoxicity and genotoxicity.[19][31] To address this issue, polymorphic UGT1A7 alleles were expressed and characterised for differences in catalytic activity. We demonstrated that glucuronidation of benzo(α)pyrene metabolites and PhIP were markedly altered (table 2). For UGT1A7*3 that was identified as a main allelic association and risk factor in CRC, catalytic analysis indicated loss of activity for all tested substrates. The UGT1A7*4 allele exhibited substantially reduced phenolic UGT activities and undetectable activity towards carcinogen metabolites such as 8- and 9-hydroxybenzo(α)pyrene and the human precarcinogen PhIP. Expression of UGT1A7*2 demonstrated substantially reduced overall UGT activity with preservation of carcinogen metabolite glucuronidation (table 2). Preservation of this activity may explain the high incidence of N129K/R131K in the normal population represented in the UGT1A7*2/*1 (21%) and UGT1A7*2/*2 (16%) alleles. The lowest glucuronidation activity was found with the UGT1A7*3 allele, which confirms previous data.[25] In an earlier study, UGT1A7 activities were analysed using stably transfected HEK cell clones in addition to a higher concentration of UDP-glucuronic acid in the UGT enzymatic assays. In our assay, UGT1A7*3 was characterised by undetectable UGT activity with the substrates shown in table 2, which was confirmed by several individual transfection experiments. Although Guillemette *et al* showed some remaining UGT activity of UGT1A7*3 in their stably expressing clones, both studies arrived at the conclusion that UGT1A7*3 represented the allele with the least benzo(α)pyrene metabolite glucuronidation activity. The catalytic analysis supports the finding that the UGT1A7*3 allele is significantly associated with the cancer group, identifying it as the main risk allele of CRC.

The magnitude of the altered catalytic activities encoded by the polymorphic alleles is an unexpected finding based on structural and biological considerations. The UGT1A7*3 allele has three amino acid changes, all of which were conserved with either UGT1A8 or UGT1A9 (fig 2). The high degree of homology between UGT1A7, UGT1A8, UGT1A9, and UGT1A10 has led to the assumption of great similarity of substrate specificity, as evidenced by glucuronidation of many phenolic compounds.[3][4] However, examined in the absence of homology to other UGTs, these changes could be perceived to

Table 2 Alteration in catalytic activity: UGT1A7 polymorphisms (activities in pmol/min/mg)

Substrate	UGT1A7*1 wild-type	UGT1A7*4 W208R	UGT1A7*2 N129K/R131K	UGT1A7*3 N129K/R131K/W208R
1-Nitrophenol	1773 (615)	32 (23)	13 (2)	nd
1-Naphthol	592 (100)	36 (16)	nd	nd
4-Methylumbelliferone	12110 (361)	191 (136)	13 (2)	nd
4-tert-Butylphenol	121 (2)	nd	12 (0.2)	nd
7-Hydroxybenzo(α)pyrene	1263 (20)	51 (10)	14 (0.5)	nd
8-Hydroxybenzo(α)pyrene	98 (6)	nd	12 (1)	nd
7–8-Hydroxybenzo(α)pyrene	79 (7)	nd	15 (1)	nd
9-Hydroxybenzo(α)pyrene	198 (15)	nd	16 (3)	nd
PhIP	48 (4)	nd	17 (5)	nd

Mean catalytic activities of the missense mutations identified in exon 1 of UGT1A7. The wild-type UGT1A7*1 exhibited high activities with phenolic compounds as well as with benzo(α)pyrene metabolites, as previously characterised.[3] The analysis demonstrates reduced glucuronidation activities of the polymorphic UGT1A7 proteins. The major risk allele identified in the patient analysis, UGT1A7*3, exhibited the lowest carcinogen detoxification activity.
PhIP, amino-1-methyl-6-phenylimidazo-(4.5-β)-pyridine (PhIP); nd, not detected.
Values are mean (SD).

impact on function as the N129K and W208R differences represent non-conservative substitutions.

The majority of CRC patients with the UGT1A7*3 (129K/131K/208R) genotype were heterozygous (table 1). Similar findings have been reported for the mismatch repair genes in familial CRC.[13][16] In the *UGT1A* genes, the presence of heterozygous mutant alleles with a reduced activity towards carcinogens can be influenced by additional occurrence of polymorphisms in exons 2–5 which are common to all UGT1A gene products. This combination would introduce an added risk factor capable of modifying the catalytic activity of all UGT1A proteins and would exert an increased influence on the already altered activity of polymorphic *UGT1A7* alleles described in this study. In view of the previous identification of 20 individual polymorphisms located within the common UGT1A exon 2–5 sequences (reviewed by Tukey and Strassburg[4]), additional combinations with the *UGT1A7* polymorphisms described here are likely.[32]

Our findings implicate the combination of genetic predisposition and environmental factors via the cellular detoxifying enzyme UGT1A7 in human CRC. The UGT1A7*3 allele exhibited a highly significant association with CRC. It was present in only 27% of the normal population studied (allelic frequency 0.17) but a tumour free state at the time of blood sampling does not preclude the development of CRC in normal controls at a later time. This suggests that screening for the presence of UGT1A7*3 would have a negative predictive value of 73%. Screening for 129K/131K/208R of the human *UGT1A7* gene in selected patients may represent an attractive diagnostic tool to identify interindividual CRC risk profiles. Based on these findings, sporadic CRC may have a higher degree of individual genetic predisposition than previously thought. As CRC is a potentially curable disease when diagnosed at an early stage, the clinical implementation of our findings has the potential to improve the primary prevention of CRC in risk groups and to affect the decision process for endoscopic evaluation. In addition, *UGT1A7* may serve as an attractive candidate gene for gene therapeutic strategies.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations
C P Strassburg, A Vogel, S Kneip, M P Manns, Department of Gastroenterology and Hepatology, Hannover Medical School, 30625 Hannover, Germany
R H Tukey, Departments of Chemistry and Biochemistry, and Pharmacology, University of California, San Diego, La Jolla, CA 92093, USA

REFERENCES

1 **Strassburg CP**, Oldhafer K, Manns MP, *et al.* Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. *Mol Pharmacol* 1997;**52**:212–20.

2 **Strassburg CP**, Nguyen N, Manns MP, *et al.* Polymorphic expression of the UDP-glucuronosyltransferase UGT1A gene locus in human gastric epithelium. *Mol Pharmacol* 1998;**54**:647–54.

3 **Strassburg CP**, Manns MP, Tukey RH. Expression of the UDP-glucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8. *J Biol Chem* 1998;**273**:8719–26.

4 **Tukey RH**, Strassburg CP. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu Rev Pharmacol Toxicol* 2000;**40**:581–616.

5 **Strassburg CP**, Kneip S, Topp J, *et al.* Polymorphic gene expression and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine. *J Biol Chem* 2000;**275**:36164–71.

6 **Strassburg CP**, Strassburg A, Nguyen N, *et al.* Regulation and function of family 1 and family 2 UDP-glucuronosyltransferase genes (UGT1A, UGT2B) in human oesophagus. *Biochem J* 1999;**338**:489–98.

7 **Strassburg CP**, Nguyen N, Manns MP, *et al.* UDP-glucuronosyltransferase activity in human liver and colon. *Gastroenterology* 1999;**116**:149–60.

8 **Strassburg CP**, Manns MP, Tukey RH. Differential down-regulation of the UDP-glucuronosyltransferase 1A locus is an early event in human liver and biliary cancer. *Cancer Res* 1997;**57**:2979–85.

9 **Parkin DM**, Pisani P, Ferlay J. Global cancer statistics. *CA Cancer J Clin* 1999;**49**:33–64.

10 **Peel DJ**, Ziogas A, Fox EA, *et al.* Characterization of hereditary nonpolyposis colorectal cancer families from a population-based series of cases. *J Natl Cancer Inst* 2000;**92**:1517–22.

11 **Das Gupta R**, Kolodner RD. Novel dominant mutations in Saccharomyces cerevisiae MSH6. *Nat Genet* 2000;**24**:53–6.

12 **Fishel R**, Kolodner RD. Identification of mismatch repair genes and their role in the development of cancer. *Curr Opin Genet Dev* 1995;**5**:382–95.

13 **Liu B**, Parsons R, Papadopoulos N, *et al.* Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat Med* 1996;**2**:169–74.

14 **Kolodner RD**. Mismatch repair: mechanisms and relationship to cancer susceptibility. *Trends Biochem Sci* 1995;**20**:397–401.

15 **Dunlop MG**, Farrington SM, Carothers AD, *et al.* Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum Mol Genet* 1997;**6**:105–10.

16 **Lynch HT**, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* 1999;**36**:801–18.

17 **Kadlubar FF**, Butler MA, Kaderlik KR, *et al.* Polymorphisms for aromatic amine metabolism in humans: relevance for human carcinogenesis. *Environ Health Perspect* 1992;**98**:69–74.

18 **Nicolaides NC**, Littman SJ, Modrich P, *et al.* A naturally occurring hPMS2 mutation can confer a dominant negative mutator phenotype. *Mol Cell Biol* 1998;**18**:1635–41.

19 **Hursting SD**, Slaga TJ, Fischer SM, *et al.* Mechanism-based cancer prevention approaches: targets, examples, and the use of transgenic mice. *J Natl Cancer Inst* 1999;**91**:215–25.

20 **Vasen HF**, Watson P, Mecklin JP, *et al.* New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* 1999;**116**:1453–6.

21 **Bansal SK**, Gessner T. A unified method for the assay of uridine diphosphoglucuronyltransferase activities toward various aglycones using uridine diphospho[U-14C]glucuronic acid. *Anal Biochem* 1980;**109**:321–9.

22 **Bartsch H**, Petruzzelli S, De Flora S, *et al.* Carcinogen metabolism in human lung tissues and the effect of tobacco smoking: results from a case-control multicenter study on lung cancer patients. *Environ Health Perspect* 1992;**98**:119–24.

23 **Guengerich FP**. Roles of cytochrome P-450 enzymes in chemical carcinogenesis and cancer chemotherapy. *Cancer Res* 1988;**48**:2946–54.

24 **Guengerich FP**, Shimada T, Iwasaki M, *et al.* Activation of carcinogens by human liver cytochromes P-450. *Basic Life Sci* 1990;**53**:381–96.

25 **Guillemette C**, Ritter JK, Auyeung DJ, *et al.* Structural heterogeneity at the UDP-glucuronosyltransferase 1 locus: functional consequences of three novel missense mutations in the human UGT1A7 gene. *Pharmacogenetics* 2000;**10**:629–44.

26 **von Heijne G**. Signals for protein targeting into and across membranes. *Subcell Biochem* 1994;**22**:1–19.

27 **Martin L**, Assem M, Piard F. Are there several types of colorectal carcinomas? Correlations with genetic data. *Eur J Cancer Prev* 1999;**8**(suppl 1):S13–20.

28 **Hill MJ**. Molecular and clinical risk markers in colon cancer trials. *Eur J Cancer* 2000;**36**:1288–91.

29 **Gertig DM**, Stampfer M, Haiman C, *et al.* Glutathione S-transferase GSTM1 and GSTT1 polymorphisms and colorectal cancer risk: a prospective study. *Cancer Epidemiol Biomarkers Prev* 1998;**7**:1001–5.

30 **Brockton N**, Little J, Sharp L, *et al.* N-acetyltransferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol* 2000;**151**:846–61.

31 **Mayhew JW**, Lombardi PE, Fawaz K, *et al.* Increased oxidation of a chemical carcinogen, benzo(a)pyrene, by colon tissue biopsy specimens from patients with ulcerative colitis. *Gastroenterology* 1983;**85**:328–34.

32 **Strassburg CP**, Manns MP. Jaundice, genes and promoters. *J Hepatology* 2000;**33**:476–9.