Deduced amino acid sequence, gene structure and chromosomal location of a novel human Class Mu glutathione S-transferase, GSTM4

Shan ZHONG,* Nigel K. SPURR,† John D. HAYES‡ and C. Roland WOLF*§

*ICRF Molecular Pharmacology Group, Hugh Robson Building, George Square, Edinburgh EH8 9XD, Scotland, †ICRF Human Genetic Resources Laboratory, Clare Hall Laboratories, Blanche Lane, Potters Bar, Herts. EN6 3LD, and ‡University Department of Clinical Biochemistry, Royal Infirmary, Edinburgh EH3 9YW, Scotland, U.K.

The Mu-Class glutathione *S*-transferases (GSTs) are subject to marked inter-individual variation in man, owing to the fact that 40–50 % of the population fail to express M1 subunits. Mu-Class GST from two lymphoblastoid cell lines (expressing M1 subunits and the other 'nulled' for M1) have been studied. Both cell lines were found to express a Mu-Class GST that has not been described previously. The cDNA encoding this novel transferase, designated 'GSTM4', has been isolated and the enzyme shown to be comprised of 218 amino acids (including the initiator methionine residue) with an M_r of approx. 25.5 kDa. Molecular cloning demonstrated that the lymphoblastoid cell line which expressed GSTM1 possessed the b allelic variant (i.e. that with an asparagine residue at position 173). The genes for GSTM4 and GSTM1b have been cloned and found to contain seven introns and eight exons. The coding region of the GSTM4 gene, including the seven introns, encompasses 5.0 kb, whereas the same region of GSTM1b is 5.5 kb; the difference in the size of the two genes is due to the length of intron 7. DNA sequencing allowed a GSTM4-gene-specific oligo-primer to be designed which has been utilized in a PCR-based assay to determine that the GSTM4 gene is located on chromosome 1.

INTRODUCTION

The glutathione S-transferases (GST), through their ability to catalyse the conjugation of GSH with electrophilic xenobiotics, provide protection against carcinogens, environmental pollutants and chemotherapeutic drugs (Chasseaud, 1979; Hayes et al., 1990). Such compounds include insecticides, herbicides, antitumour agents and epoxide-containing metabolites of polycyclic aromatic hydrocarbons (Hayes and Wolf, 1988).

Mammalian GST comprise five families. Four of these are cytosolic and have been designated Alpha, Mu, Pi and Theta, whereas the fifth is membrane-bound and is simply referred to as 'microsomal GST' (Mannervik et al., 1992). Within each gene family the GST share at least 55 % amino acid sequence identity, but between families identity is less than 25 %. Present evidence suggests that the Alpha and Mu families are the most complex, both possibly comprising at least six genes, whereas the Pi and microsomal families comprise only one or two genes (Board et al., 1990). Further complexity in the number of isoenzymes arises because GSTs are dimeric and because subunits that are members of the same family can heterodimerize (Hayes et al., 1981; Stockman et al., 1985). A large number of heterodimers are found in most tissues (Hussey et al., 1991).

In man, the GSTs which comprise the Mu evolutionary class have attracted particular interest because they are subject to marked inter-individual differences (Warholm et al., 1980; Board, 1981; Strange et al., 1984; Hussey et al., 1986). Zymogram analysis has shown that approx. 45 % of Europeans fail to express a transferase encoded at the *GSTM1* locus (Board, 1981; Laisney et al., 1984; Strange et al., 1984). The null phenotype is due to gene deletion (Seidegard et al., 1988). Among those individuals that express GSTM1, two common allelic variants, GSTM1*a and GSTM1*b, exist. These alleles encode GSTM1a-1a and GSTM1b-1b, previously celled μ and ψ respectively, which are functionally identical (Widersten et al., 1991), but differ in a single amino acid; GST μ contains lysine at position no. 172 (De Jong et al., 1988), whereas GST ψ contains asparagine at this position (Seidegard et al., 1988).

Since Mu-Class GSTs are particularly effective at deactivating mutagenic and carcinogenic epoxides (e.g. styrene oxide, *trans*-stilbene oxide, 1-nitropyrene oxide, benzo[*a*]pyrene-7,8-diol 9,10-epoxide), it is reasonable to postulate that the polymorphic expression of this family results in a significant reduction in the detoxification capacity of nulled individuals. Seidegard et al. (1986, 1990), who monitored expression at the *GSTM1* locus by measuring leucocyte transferase activity towards *trans*-stilbene oxide, reported that failure to express GSTM1 is associated with an increased susceptibility to lung cancer in cigarette smokers. In a subsequent study we have found a small positive correlation between squamous carcinoma of the lung and the null genotype (Zhong et al., 1991). Surprisingly, a negative correlation between the null genotype and adenocarcinoma of the lung was also observed. Strange et al. (1991), using a starch-gel zymogram method, have found an increased frequency of the null phenotype in patients with adenocarcinoma of the stomach or colon.

Much remains to be learnt about the molecular genetics of the human Mu-Class GSTs. Besides the molecular cloning of GSTM1a and GSTM1b, two further cDNAs have been reported that encode the 'muscle-specific' transferase, GSTM2 (Vorachek et al., 1991) and a 'brain/testis-specific' transferase, GSTM3 (Campbell et al., 1990). However, it is certain that additional Mu-Class GSTs exist, since zymogram analysis of human tissue extracts suggests that this family is encoded at more than three loci (Suzuki et al., 1987), and sequence analysis of a genomic clone (COS H1-10), from a human cosmid library has revealed a Mu-Class gene distinct from those encoding GSTM1, GSTM2 or GSTM3 (Taylor et al., 1990). The cosmid clone isolated by Taylor et al. (1990) contains partial sequences from two Mu-

Abbreviations used: GST, glutathione *S*-transferase; DMSO, dimethyl sulphoxide.

§ To whom correspondence should be sent.

The nucleotide sequences reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers GSTM16, X68676, GSTM4 and X68677.

Class GST genes that these workers designated 'GSTmu2' and 'GSTmu3'; amino acid sequencing of the muscle-specific transferase suggests that GSTmu3 is equivalent to GSTM2 (see Hussey et al., 1991), but the identity of GSTmu2 is uncertain.

Human Mu-Class genes have been reported to be located on both chromosome 1 and chromosome 3. De Jong et al. (1988) carried out *in situ* hybridization using a cDNA for GSTM1a–1a and showed that human Mu-Class genes were clustered in region p3.1 of chromosome 1. This assignment was confirmed by PCR mapping (Zhong et al., 1993). By contrast, using a panel of human × rodent somatic-cell hybrids and two different rat Mu-Class cDNA probes, Islam et al. (1989) assigned Mu-Class genes to chromosome 3. The second group of workers postulated that they had mapped GSTM2 and GSTM3. However, as Taylor et al. (1990) have found two separate Mu-Class genes within a single cosmid clone, it remains unclear whether all the Mu-Class GST genes are on the same chromosome.

Further molecular information about Mu-Class GST genes is required to allow a better understanding of the multiplicity of this family in Man and their involvement in protection against carcinogenesis. To this end, we describe the isolation of a cDNA encoding human GSTM4. The gene structure of GSTM4 has been determined, and its chromosomal localization indicates that at least three of the characterized GSTM genes are found on chromosome 1.

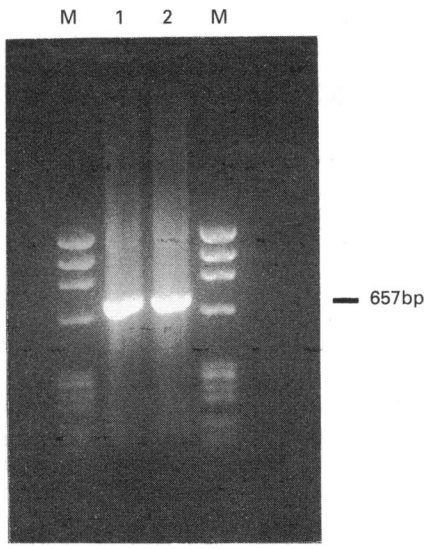

Figure 1 Amplification of GSTM1b and GSTM4 cDNAs from lymphoblastoid cell lines

M, phage ϕX174/*Hae*III marker fragments; lanes 1 and 2, the 657 bp DNA fragment from GSTM1 cDNA from a GSTM1-positive cell line and GSTM4 cDNA from a GSTM1 null cell line respectively.

MATERIALS AND METHODS

Nomenclature

The nomenclature used in this study is that recommended for GST by Mannervik et al. (1992).

Cell lines and preparation of DNA and RNA

The lymphoblastoid cell lines GST5 and GST6 were obtained by Epstein–Barr-virus immortalization of normal human lymphocytes. Cells were grown in RPMI 1640 and 10% foetal-calf serum. These lines have been phenotyped and genotyped for hGSTM1 (Zhong et al., 1991), and the somatic hybrids have been described previously (Zhong et al., 1993). The human chromosome content of the hybrids were checked by a combination of isoenzyme and DNA marker analysis and karyotypic analysis (Evans et al., 1971; Bobrow and Cross, 1974; Harris and Hopkinson, 1976). Genomic DNA was extracted from those cell lines and appropriate controls by using an Applied biosystems DNA extractor 340A, the manufacturer's recommendations being followed. Total cellular RNA was prepared from lymphoblastoid cell lines by the guanidine thiocyanate method (Chirgwin et al., 1979). About 10 μg of RNA was mixed with 600 ng of primer P4 (complementary to the 3'-end of hGSTM1 cDNA) for full-length cDNA synthesis using the cDNA synthesis kit.

Synthesis of oligonucleotide primers

The following oligonucleotide primers were prepared on an Applied Biosystems 370 oligonucleotide synthesizer:

These primers were used in separate PCRs to generate the hGSTM4 and hGSTM1b cDNAs and genes. The positions of the above primers are indicated in Figure 1. The additional primers for sequencing were synthesized as sequences were obtained.

PCR

PCR was carried out in a total volume of 100 μl. For amplification of genomic DNA the following reaction mixtures were employed: 1 μg of DNA, 600 ng of each primer, 300 μM of dNTPs (Pharmacia), 10 μl of DMSO (dimethyl sulphoxide), 10 μl of 10 × PCR (1 × PCR buffer is 10 mM Tris/HCl, pH 8.3, containing 50 mM KCl, 1.5 mM $MgCl_2$ and 0.001% gelatin) buffer (containing 15 mM $MgCl_2$) and 2 units of *Taq* DNA polymerase (Promega). The sample was covered with 100 μl of light mineral oil (Sigma). The PCR was allowed to proceed for 30 cycles (each consisting of 1 min at 94 °C, 2 min at 55 °C and 3 min at 72 °C) which was followed by a 10 min incubation at 72 °C using a programmable thermocycler (M.J. Research). For amplification of cDNA the concentration of dNTPs was reduced to 100 μM and no DMSO was added. The PCR product was analysed by electrophoresis through an agarose gel in TAE buffer (0.04 M Tris/acetate, 0.001 M EDTA, pH 8.3). The gel was stained in ethidium bromide and photographed.

Cloning and sequencing

A 20–50 μl aliquot of the product of the PCR reaction was run on a 1% low-melting-point agarose gel and the desired band was removed from the gel under the long-wave u.v. light. The purified DNA fragment was ligated into the PCR-1000 vector (Invitrogen) using T4 ligase (Pharmacia) then electrotransfected to the

P1:	5'-ATGCCCATGATACTGGGGTACTGG-3'	(sense)
P1:	5'-GTGGAGGTCAAGGACATCATAG-3'	(antisense)
P2(M1):	5'-GGAATGAGATCTGTTTTGCTTCAC-3'	(sense)
P3(M4):	5'-CACTAGAGTTTTGCCATACATCCT-3'	(sense)
P4:	5'-CTACTTGTTGCCCCATACATCCAT-3'	(antisense)

```
                1        10        20        30        40        50        60        70
                |        |         |         |         |         |         |         |
hGSTM4          MPMILGYWDIRGLAHAIRLLLEYTDSSYEEKKYTMGGAPDYDRSQWLNEKFKLGLDFPNLPYLIDGAHKITQSNAILCY
hGSTM1b         .....................................D.....................................
hGSTM2          ...T....N......S.....................D.............................T.........R.
hGSTM3          MSCESS.V.....................F..T.....R..C.E..........DV....D.........L..KN.........R.
rGSTM1          ........NV...T.P..............R.A.............................SR........MR.
rGSTM2          ...T..............F.....T...D...S..D..........S.............S..........R.
rGSTM3          ...T.............................R....D...F.....................S..........R.
mGSTM1          ........NV...T.P..M.........D..R....D...F.....................S..........R.
mGSTM2          ...T.....................T...D......D..........S.............S..........R.
mGSTM3          ...T....NT...T.S..............R.V..D..NF......S...N.............S..V.......R.
gpGSTM          ...T.....N....T.P...I....N.G....R.N.............................T..L.......R.
haGSTM          ..VT.....................T..........D..NF.....................S..........R.
chGSTM          .VVT.....................ETP.QERR.KA.P...E.P.D.T.................DV.L.......R.
              <---EXON 1----><----------EXON 2--------><-------EXON 3-------><-----------EXON 4---

                80        90        100       110       120       130       140       150
                |         |         |         |         |         |         |         |
hGSTM4          IARKHNLCGETEEEKIRVDILENQAMDVSNQLARVCYSPDFEKLKPEYLEELPTMMQHFSQFLGKRPWFVGDKITFVDFL
hGSTM1b         .........................T..NHM..GMI..N.E......K......EKLKLY.E........A.N........
hGSTM2          ..........S.K.Q..E......F..SRM...KL..D............QA..E.LKLY......Q...L.........I
hGSTM3          ......M.............I...V..FRT..I.L...S.H.....Q......GQLKQ..M..W.FS..A.E.L......
rGSTM1          L....H........R..A..V...V..NRM..IML........Q...F.KTI.EK.KLY.E........A...V.Y....
rGSTM2          LG.............R....V.......TRL...M........RK......G..EK.KLY.E....Q...A.N...Y....
rGSTM3          LG.............R.........L..NRMV...L..N........G...Q..G..RLY.E........A.........I
mGSTM1          L....H.D......R..A..V...V..TRM..IML..N.....Q...F.KTI.EK.KLY.E........A...V.Y....
mGSTM2          L.............R............TRI...M.........K......G..EK.KLY.E....Q...A.N.V.Y....
mGSTM3          LG.............R....T....V..TRI..MI..C......Q...F.KAI.EK.KLY.E........A...V.Y....
gpGSTM          .........V....T..M......V..IRM..IML.......QK.A.F..GI.DK.KL.......L...A.N.L.Y....
haGSTM          ....D.........R.QL.........TRM...M.........R......G..EK.KLY.E.....S..A.....Y..T.
chGSTM          ......M.....V..Q...V...HL..LR.AF..L............A...L..GKLRQL.R...S.S......L......
                -----><-----------EXON 5-------------><-----------EXON 6-----------><-------

                160       170       180       190       200       210                 Identity
                |         |         |         |         |         |
hGSTM4          AYDVLDLHRIFEPNCLDAFPNLRDFISRFEGLEKISAYMKSSRFLQKPLYTRMAVWGNK
hGSTM1b         V.....................K.....................PR.VFSK.......          87%
hGSTM2          .....ERNQV...S........K.....................PR.VF.K.......          82%
hGSTM3          ...I..QN...D.K...E....KA.MC...A....A..LQ.DQ.CKM.INNK..Q....PVC      69%
rGSTM1          ...I..QYH....K........K..LA.....K..........Y.ST.IFSKL.Q.S..          74%
rGSTM2          ......Q......K........K..VA.....K...D....G...S..IFAK..F.NP.          78%
rGSTM3          .....ERNQV..AT............A.....K...D........PR....K..I..S.          82%
mGSTM1          ...I..QY.M...K...........LA.....K..........YIAT.IFSK..H.S..          74%
mGSTM2          ......Q......K........K..MG.....K...D........S..IFAK..F.NP.          81%
mGSTM3          ...I..QY.M...K...........LA.....K............PR.VF.KI.Q..TD          74%
gpGSTM          ......QY.ML..K..E.....K.............S........P...FSKF.F.N..          76%
haGSTM          I.....Q....A.K....A.......LA.....K...D........SC.QIFAK....NS.          79%
chGSTM          ......QQ.M.V.D.PELQG..SQ.LQ...A.........R.G..MKA.IFWYT.L.NN.KE       66%
                -------EXON 7--------------><---------EXON 8--------------->
```

Figure 2 Comparison of the deduced amino acid sequences of hGSTM4 with other GST Mu sequences

The sequences were obtained from the following sources. hGSTM1b (Seidegard et al., 1988); hGSTM2 (Vorachek et al., 1991); hGSTM3 (Campbell et al., 1990); rGSTM1 (Lai et al., 1986); rGSTM2 (Lai et al., 1986) rGSTM3 (Abramovitz and Listowsky, 1987); mGSTM1 (Pearson et al., 1988); mGSTM2 (Townsend et al., 1989); mGSTM3 (Pearson et al., 1983); gpGSTM (Kamei et al., 1990); and chGSTM (Liu et al., 1991). Only amino acid differences with respect to hGSTM4 are shown. The sequence identity of GSTM4 with the other GSTs is given in the last column. The prefixes are defined as follows: designate: h, human; r, rat; m, mouse; gp, guinea-pig; ha, hamster; Ch, Chinese hamster.

Escherichia coli host INVaF by the Electroporator (Invitrogen). Single recombinant PCR-1000 colonies were grown up to prepare double-stranded plasmid DNA using Qiagen columns (Hybaid). The DNA sequencing was carried out using the method of Sanger et al. (1977), using a Sequenase kit (U.S. Biochemicals) and [α-^{35}S]thio-dATP (Amersham), as modified by Hsiao (1991). Both strands of DNA were sequenced. Sequence data was computer-analysed by using the IntelliGenetics (IntelliGenetics) and Gene Jockey (Biosoft) software packages.

RESULTS AND DISCUSSION

Molecular cloning of cDNA encoding human Mu-Class GST

The Western blotting and genotype assays developed previously (Seidegard et al., 1988; Zhong et al., 1991) allow lymphoblastoid cell lines which express GSTM1 (a or b) to be distinguished from those that are nulled at the GSTM1 locus (i.e. GSTM1*0/GSTM1*0). Total mRNA was extracted from a lymphoblastoid cell line containing the GSTM1 gene, and the cDNAs encoding Mu-Class GSTs were selectively reverse-transcribed using the P4 oligonucleotide to prime the reaction. The cDNAs were then amplified by PCR from the oligo-primers P1 and P4, and analysis of the PCR products by agarose-gel electrophoresis revealed a single band of approx. 650 bp. Surprisingly, amplification of Mu-Class GST cDNA from a lymphoblastoid cell line nulled for GSTM1, which was employed as a negative control, yielded a band of identical electrophoretic mobility (Figure 1). The PCR products obtained from the cell line expressing GSTM1 (a or b) were cloned into the PCR-1000 vector, and five separate colonies were selected for DNA sequencing. Three of the five clones had identical GSTM1 sequences and shared complete identity with the human Mu-Class GST cDNA encoding transferase ψ (hGSTM1b) reported by Seidegard et al. (1988), i.e. with asparagine at position 173 (including the initiator methionine) rather than the lysine that occurs in the case of GSTM1a (see De Jong et al., 1988). The sequences of the inserts of the remaining two clones were identical with each other, but were distinct from previously described GST cDNAs. This novel nucleotide sequence has an open reading frame of 654 bp that encodes a protein with a molecular mass of about 25.5 kDa. The deduced polypeptide shares 72–88% sequence identity with other members of the Mu-Class GST family, and, as it is distinct from GSTM1, GSTM2 and GSTM3, it has been designated 'GSTM4' (Figure 2).

The PCR product obtained from the lymphoblastoid cell line possessing the GSTM1*0/GSTM1*0 genotype was also cloned, and four clones containing the insert were sequenced. These were all identical with the GSTM4 cDNA isolated from the GSTM1b-expressing cell line. Therefore the sequence for hGSTM4 has been deduced from a total of six clones which were obtained from two independent lymphoblastoid cell lines following separate PCR reactions.

Among the human enzymes, GSTM4 is more closely related to GSTM1 (a/b) and GSTM2 than to GSTM3 (Figure 2). At the protein level, human GSTM4 also possesses a high degree of sequence identity (approx. 78–82%) with rat Yb_2 (rGSTM2), rat Yb_3 (rGSTM3) and mouse Yb_2 (mGSTM2) subunits. Interestingly, the majority of the differences in primary structure between these GST subunits and GSTM4 are clustered between residues 103 and 116, 132 and 139, 165 and 172, and 204 and 210 (the numbering of residues excludes the initiating methionine).

Organization and sequence of hGSTM1b and hGSTM4 genes

The human GSTM1b and GSTM4 genes were amplified by PCR using genomic DNA from the two lymphoblastoid cell lines described above. Initial attempts to amplify the entire GSTM1 and M4 genes using the oligonucleotides P1 and P4 were unsuccessful. However, these genes were isolated by amplifying two overlapping gene fragments in separate PCR reactions. The region between exon 1 and exon 7 in both GSTM1b and M4 were amplified using oligonucleotide primers P1 and P2, corresponding to conserved sequences between these genes. The regions between intron 6 and exon 8 in GSTM1b and GSTM4 were selectively amplified using gene-specific primers P3 (M1) (Figure 4a below) or P3(M4) (Figure 4b below) for intron 6 along with P4. An oligonucleotide corresponding to the conserved C-terminal portion of Mu-Class GST (the P3 gene-specific oligo-primers for intron 6) could only be designed once the PCR products generated from P1 and P2 were sequenced.

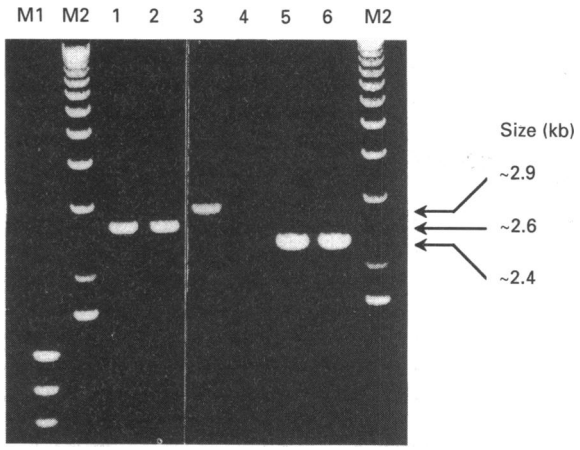

Figure 3 Analysis of hGSTM1b and hGSTM4 genes

Genomic DNA was amplified by PCR using the oligonucleotide primers P1 and P2 or P3(M1), P3(M4) together with P4 (Figure 4). M1, phage φX174/*Hae*III marker fragments; M2, 1 kb ladder marker fragments; the DNA fragments shown in lanes 1, 3 and 5 were obtained from genomic DNA containing the hGSTM1b and hGSTM4 genes. Lanes 2, 4, and 6 show the DNA fragments obtained from a GSTM1 nulled cell line. These were shown to be derived from the hGSTM4 gene. The fragments were obtained: lanes 1 and 2 using primers P1 and P2; lanes 3 and 4, using hGSTM1b-gene-specific primer P3(M1) and P4; lanes 5 and 6, using hGSTM4-gene-specific primer P3(M4) and P4.

Genomic DNA extracted from the cell line nulled for GSTM1 gave a PCR product of 2.6 kb using primers P1 and P2 (track 2, Figure 3), whereas primers P3(M4) and P4 produced a single DNA fragment of 2.4 kb (track 6). Three separate clones from each PCR were sequenced, and they unambiguously showed that these two PCR fragments contained exons 1–8 of the hGSTM4 gene and include a 90 bp overlap. The hGSTM1b gene was cloned from the lymphoblastoid cell line expressing GSTM1b using a combination of primers P1 and P2 or P3(M1) and P4. The PCR product from P3(M1) and P4 was 2.9 kb in length compared with 2.4 kb in the case of GSTM4 (Figure 3, track 3). Three clones which possessed the 2.9 kb insert were sequenced and found to contain exon 1–exon 8 from hGSTM1b. Three clones from separate PCR reactions were sequenced and confirmed that the cloned PCR fragments were derived from the hGSTM1b gene. Electrophoretic analysis suggests that exon 1–exon 8 of GSTM1b is 5.5 kb in length, whereas the same portion of GSTM4 is only 5.0 kb (Figure 3).

Sequence analyses revealed that the organization of hGSTM1b and hGSTM4 genes is similar to that of the rGSTM2 gene (rat Mu-Class GST Yb_2 subunit gene) described by Lai et al. (1988). These genes all comprise eight exons that are interrupted by seven introns, and splicing obeys the gt/ag splice rule. The genomic coding sequences of the hGSTM1b and hGSTM4 genes

Figure 4 **For legend see page 47**

4401 caggtagtcgagtggtttctaagcttatgttgtaatttctcttggtacagagcacccagcaccgtgtagaatcttcataa
4481 gtgttagatgttactgtggtacaacattacttaaggaagttggaagagttaactccgcaaatctggggaccctaagaggc
4561 tgtgtgatgcctcagcacttgagcccacgtggaaaggctgtgccagggccctgacctgctgtgtctgcagtggggttgtc
4641 ccacgccgtgaagctgtgacagggaccctgacctgctgtgtctgcagtggggttgtcccaagcctcatgggcagctgacc
4721 ttgagttctggccttatttccccctctcag**GGCTTGGAGAAGATCTCTGCCTACATGAAGTCCAGCCGCTTCCTCCCA**

P4

4801 **AAACCTCTGTACACAAGGATGGCTGTCTGGGGCAACAAGTAG**

(b)

P1

1 **ATGCCCATGATACTGGGGTACTGGGACATCCGTGGG**gtaagcgagggtcctctggtgggtgggacaggggcggaggcgg
81 ggatgtgtggagtagctgcaggactggctctagggaccgttcctcttcagggctgcccgcctcagaagggcctgtgcatg
161 acgctgtgtgtgtgtttgggggtgggggcgggtagaggaggcgacgggtacgtgcagtatagactagggctggcctggtg
241 cagagaaagtcaccaagtcagggaccctccatctctgacccgagccgcggccatctctcccag**CTGGCCCACGCCATCCG**
321 **CCTGCTCCTGGAATACACAGACTCAAGCTACGAGGAAAAGAAGTATACGATGGGGGACG**gtaatgacacccttgtgtccg
401 ggctctgcactcacgctgagttggcaccaagcaacccatggtggccacctgtcgtacctctgcaggcctcccctgctgga
481 gctgcaggctgtcccttccctgagccccggtgaggagtcctgtggccttgcaaggcagaatgctggggcgggatagtggg
561 tccctgtttaattgggttgggtgtcctcagagcttcccaaaccctggaagccttagccgtgtggggtccagagcctcagc
641 gggattatttgtccctgaaccctgggatgtgggactgagtggtcagattctagatccacctgtctcagggatcttgccac
721 tggttccttgggagggtccccggaaggagggctgggctctggggaggtttgttttcacttcttccccacggcag**CTC**
801 **CTGACTATGACAGAACGCAGTGGCTGAATGAAAAATTCAAGCTGGGCCTGGACTTTCCCAAT**gtaggtgcagggggaagg
881 ggcggtttgggggaaagtgcgacgtgtctctgactgcatctcctctccccagattagaggtgttcggatcaggagtctt
961 ctgcccaattcctggttgtctacacagcccctgcatgatgttctgtgtcccagctcatttgttcatgtgacagtatttct
1041 atgtcaggcctgcatgagcgggcacagtgagtctggtctccccttgcatataggaaggggatgctggggagcctgctggc
1121 cccaactgagcttccccggtttcccatctatccag**CTGCCCTACTTGATTGATGGGGCTCACAAGATCACCCAGAGCAAC**
1201 **GCCATCTTGTGCTACATTGCCCGCAAGCACAACCTGT**gtgagtgtgggtggctgcaatgtgtgggggaaggtggcctcc
1281 tccttggctgggctgtgatgctgagattgagtctgtgttttgtgggtggcag**GTGGGGAGACAGAAGAGGAGAAGATTCG**
1361 **TGTGGACATTTTGGAGAACCAGACCATGGACAACCATATGCAGCTGGGCATGATCTGCTACAATCCAGAATTT**gtgagtg
1441 tccccagtgagctgcatctgacagagtttggatttggggccaggactcttgcatcctgcacacattggtcttaagtccct
1521 ggtaccattcatcctccaagtgctttcccatcatctagcagtatctctacgactccaatgtcatgtcaacaaaagcagag
1601 gcaattcccaaccaaccttaggacacgattccaggcattcccagggtagaaatttcagttcctgtatggtaaagtttgtg
1681 ttcagaatctccttcatcagctctggcctctgacttctgtcctgggtcatttctgtcagccagttcacatcacctgcctg
1761 ctcctagaatatgcagactcaagtagaagactcaggaatgtaatggcaccctcgaattgcatcttctcctcaacagtttt
1841 ctgagtgctgtcattgacatgcacagggatctgcgcatcttcataacagacagctcagaggcagtcagagggcctttatt
1921 cctctccctccttcctttcaacttgaacttctcatctccctggaaactagtcaacgttcattgtttcttctgccacccc
2001 attagaaggaactttctactttccctgagctcccttagttctttgcatccttgattctgctggtctggatccagaggctg
2081 ccaggtgcttgggcgctcctggggctgacccagaggctattgggaggtcagtgaggacagattcagggacagcatctcat
2161 tcctctctgccttctgatcagtttagatagggtctgacactcagtcagagtctaaaatgctgagtatccaattgaagcct
2241 gcactgccccagttccagacttggggaagatggctgcttgcccgtgccagcctggccgtccacagccccggggaggccac
2321 gtctgtgcagggagcttttgtccgagggtggtgacagctgttttctgcctcag**GAGAAACTGAAGCCAAAGTACTTGGAG**
2401 **GAACTCCCTGAAAAGCTAAAGCTCTACTCAGAGTTTCTGGGGAAGCGGCCATGGTTTGCAGGAAACAAG**gtaaaggagga

P3 (M1)

2481 gtgatatggggaatgagatctgttttgcttcacgtgttatggaggttccagcccacacattcttggccttctgcag**ATCA**

P2

2561 **CTTTTGTAGATTTTCTCGTCTATGATGTCCTTGACCTCCACCGTATATTTGAGCCCAACTGCTTGGACGCCTTCCCAAAT**
2641 **CTGAAGGACTTCATCTCCCGCTTTGAG**gtgatgcccccatcctccttctctttgatgccccttgttccgttacctccttt
2721 cagatgctttcccagtcctggagctacacacagaataactcgcatgtattgagtactggtttcatgccacgaaccgtacc
2801 ccagcacattatacctattgtgtgaaatttgaattttataacattccagtaaggtaacagaattatctcgcccattttag
2881 agataaggaaactaagaatgagagggtcggtcctctgctcagggtcccagagctagtggaggcagtgctgggcccctgtg
2961 agcctctggatctatgggtggcagtcaggctctcccattcgacagagaaaagccttagcgttcacctagcctgggtttc
3041 acagcccaggacactttggaagaggcagagaacttcatgaccatagatggagctggcaatagtaggactgacacaacggt
3121 gacattgatgtctagtactgaacccacaggcaatctcatagctacctccagaagctttgcatgattggaccccagtgtgg
3201 gaatcctgagagccagggctgtggctgtagctggattaaggtacatatgtgggtgtccctgttgaaggagtatatgttga
3281 aatgcccggtgctggggcacttacttactccaccactatcttttttttttttttttttttttttgtgctggagtcttg
3361 ctctgttgcccaggctggagttcaatggagtgatcttggctcactgcaacctccgcctcctgggttcaagcgattctact
3441 gcctcagctgcacgattagttgggattacaggtgtgcaccaccacgtctggctaattttgtattttagtagagatggg
3521 gttttgccatgttggtcaggctggtctcgaactcctgacctcaggtgatctacccacatcagcctccctc--~0.6kb-tg

Figure 4 For legend see facing page

~4201 agatcgtgtcttgctgttgcccaggctggagcagcagttgcgtgacctcggacttactgcaacctctgctcccgggttca
~4281 aacaattctctgcctcagcctcccgagtagctgggaattacaagtgtctatcaccacgcccagctaatttttctatttttt
~4361 agtagagatgggcttttcaccatgttggccaggtggtcttgaactcctgacctcggtgatccacccacctcggcttccca
~4441 atgtgctgggtatacaggtgtgagccggatgtttttgaatacctatctgggcattcaatcctggtaagattgcacagca
~4521 catctgagtgtcatgtagcctgatctgcagcagggctgtagatgccatgggttagggcacagtgagattttgctcaggta
~4601 ttagatggagaactttggactttctgctttaaggggaatgtttagagcctagtctcntttgattttcttgtgcactgcca
~4681 ccccccattccactttcatccaggttttactgagacattgggtgagtgtgttcagagcccctttgttctgctgcaggtcc
~4761 cttctgtgtctctatacccagacaagccaagagcctccctgtggaaaaggagactgtttgtgcagtcaaggagtgacagg
~4841 gcctggtgtgaggggtggtggggcagaagaagaagagaatttgtcaggaagaggccagaactggagagagacagaaccag
~4921 gctacacygcaagttctattcccctacaaggtatctaaacgtaaggaagttgctgaacttctgttccacatgagaatgg
~5001 tgataatagattcagccttgcagagcagtcgagtggttttctaagcttacgttgtaatttgtgttggtacagagcaccca
~5081 gcaccgtgtagaatcttcgtaagtgttagctgttactgtggtacaacattacctaaaggaagttggaagagttaactcag
~5161 caaatctggggaccctaagaagctgtgtgatgcctcagcacttgagcccacatggaaaggctgtgccagggccctgacct
~5241 gctgtgtctgcagtggggttgtcccaccgctcatgggcagctgaccttgagttctggccttattttccccctctcag**GG**

P4

~5321 **CTTGGAGAAGATCTCTGCCTACATGAAGTCCAGCCGCTTCCTCCCAAGACCTGTGTTCTCAAAGATGGCTGTCTGGGGCA**

~5401 **ACAAGTAG**

Figure 4 Nucleotide sequences of hGSTM4 (a) and hGSTM1b (b) genes

The exons are represented by upper-case letters and the seven introns by lower-case letters. The gt–ag splice sites are underlined. The position of PCR primers are indicated by numbered arrows.

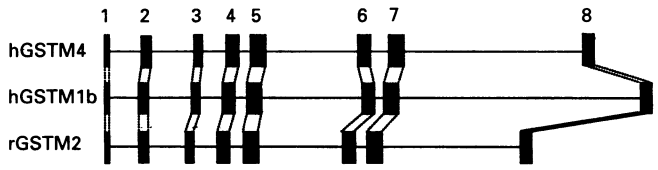

Figure 5 Comparison of the structure of the hGSTM4 gene with hGSTM1b and rGTSM2 genes

The exons are noted by numbered bars.

are identical with their respective cDNA sequences. The exons of these three Mu-Class GSTs are each identical in length and exhibit a high level of sequence similarity. By contrast, each of the introns varies significantly in size between the different Mu-Class GST genes (i.e., hGSTM1b, hGSTM4 and rGSTM2) and their sequence identity is substantially lower than that observed between the exons.

Relationship between hGSTM4 and other transferase genes

Comparison of the gene sequences of GSTM1b and GSTM4 (Figure 4) shows that introns 2, 3 and 4 share greater than 85% sequence identity, whereas introns 1 and 6 possess only 66% identity. The lowest level of sequence identity between hGSTM1b and hGSTM4 was in intron 5 (35%). The size of intron 7 differs markedly in GSTM1b and M4, which is primarily responsible for the different lengths of these two genes (Figure 3).

Table 1 Sequence similarities between the human GSTM4 gene and the human GSTM1b, GSTmu2, GSTmu3 and rat GSTM2 genes

	hGSTM4 S*	hGSTM1b S	(ID)†	rGSTM2‡ S	(ID)	GSTmu2‡ A	(ID)	GSTmu3‡ S	(ID)
Exon1	36	36	(97%)	36	(88%)				
Intron1	278	267	(66%)	268	(32%)				
Exon2	76	76	(99%)	76	(83%)				
Intron2	418	418	(85%)	319	(29%)				
Exon3	65	65	(95%)	65	(89%)	65	(98%)	65	(95%)
Intron3	306	293	(90%)	286	(27%)	310	(97%)	297	(91%)
Exon4	82	82	(97%)	82	(82%)	82	(99%)	82	(95%)
Intron4	95	95	(97%)	97	(60%)	95	(78%)	100	(89%)
Exon5	101	101	(83%)	101	(89%)	101	(100%)	101	(76%)
Intron5	915	940	(35%)	865	(27%)				
Exon6	96	96	(82%)	96	(73%)				
Intron6	90	87	(66%)	81	(44%)				
Exon7	111	111	(97%)	111	(81%)				
Intron7	2083	~2600		1784	(21%)				
Exon8	90	90	(93%)	90	(72%)				
Total	4842	~5400		4375					

* Size of nucleotides (bp).
† Identity comparison with GSTM4 sequence.
‡ The sequences for rat rGSTM2 and human GSTmu2 and GSTmu3 were taken from Lai et al. (1988) and Taylor et al. (1991) respectively.

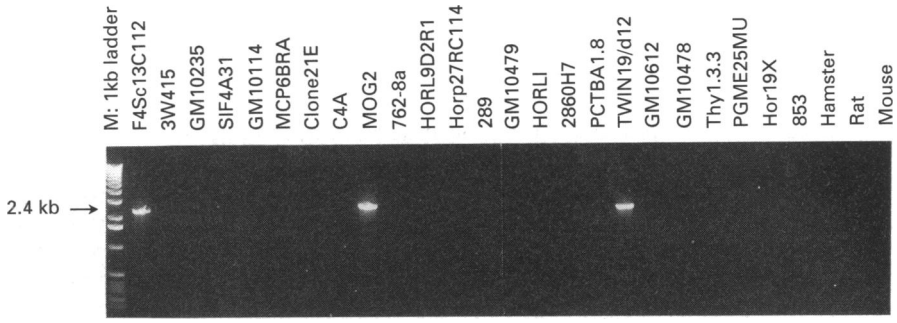

Figure 6 Mapping of the hGSTM4 gene to chromosome 1

Specific oligonucleotide primers P3(M4) and P4 were used to amplify somatic-cell hybrid DNA. The hybrids used are described in Table 2. The hybrids F4Sc13C112, MOG2 and TWIN 19/d12 all contained at least part of human chromosome 1.

The sequence identity between hGSTM4 and rat rGSTM2 is significantly lower than between the two human Mu-Class genes in all regions except exon 5 (Figure 5) (Table 1). As expected, the most substantial reduction in identity occurs in the non-coding regions, with introns 2 and 3 exhibiting the greatest divergence.

As can be seen from Table 1, hGSTM4 is distinct from other Mu-Class transferase genes described to date. However, one of the two partial Mu-Class GST gene sequences (GSTmu2) reported by Taylor et al. (1990, 1991) is highly similar to GSTM4. The 653 bp sequence of GSTmu2 described by these workers represents exon 3, intron 3, exon 4, intron 4 and exon 5. Over this region, 35 differences between GSTM4 and GSTmu2

Table 2 Assignment of the hGSTM4 locus to chromosome 1

The plus signs (+) indicate the presence of a specific human chromosome in the hybrid; +*, only a fragment of chromosome 1, i.e. 1p, is found in the cell line; t, translocation; Δ, for references about the hybrid, see Zhong et al. (1993); ΔΔ, hybrids were purchased from Human Genetic Mutant Cell Repository, National Institute of Genetic Medical Sciences (NIGMS), Camden, NJ, U.S.A. The hybrids containing the human GSTM4 gene are indicated in **bold** face (**+**) in the right-hand column.

	Human chromosomes present in human/rodent hybrid cell lines																								
Hybrid	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	**hGSTM4**	References
F4Sc13C112	+*	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−	**+**	Δ
3W4Cl5	−	+	−	−	−	−	+	−	−	+	+		−	+	+	−	+	−	−	−	+	−	+	−	−	Nabholz et al. (1969)
GM10253	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	ΔΔ
SIF4A31	−	−	+	+	−	+	−	−	−	−	−	−	−	+	−	−	+	−	−	−	−	−	+	−	−	Edwards et al. (1985)
GM10114	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	ΔΔ
MCP6BRA	−	−	−	−	−	t	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	t	−	−	Goodfellow et al. (1982)
Clone21E	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	Δ
C4A	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	Δ
MOG2	+	−	+	+	+	+	+	−	+	−	+	+	−	+	+	+	+	+	+	−	+	−	+	−	**+**	Povey et al. (1980)
762-8a	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	Δ
HORL9D2R1	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	Δ
Horp27RC14	−	+	−	+	−	−	+	−	−	−	+	+	−	+	+	−	−	−	−	−	+	+	−	−	−	Van Heyningen et al. (1975)
289	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	Δ
GM10479	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	ΔΔ
HORLI	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	Heisterkamp et al. (1982)
2860H7	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	Δ
PCTBA1.8	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	Δ
TWIN19/d12	+	−	+	−	−	+	−	+	−	−	−	+	−	+	−	+	+	+	−	+	+	+	−	−	**+**	Philips et al. (1985)
GM10612	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	ΔΔ
GM10478	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	ΔΔ
ThyB1.3	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	Goodfellow et al. (1980)
PGME25Nu	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	De Klein et al. (1982)
HORL19X	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	Heisterkamp et al. (1982)
853	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	Spurr et al. (1987)

were noted with only two changes in the exons. Most of these are within intron 4. The relationship between these two genes is at present unclear, but it is possible that they represent allelic variants.

Comparison between GSTM4 and the other partial Mu-Class GST gene sequences (GSTmu3) described by Taylor et al. (1990, 1991) shows an overall level of sequence identity of approx. 90% over the region encompassing exon 3–5. This level of identity is marginally less than that seen between GSTM4 and GSTM1b over the same region (i.e. 92% identity). If it is assumed that GSTmu3 and GSTM2 are identical (see Hussey et al., 1991), these sequencing data indicate that GSTM4 is more closely related to GSTM1 than GSTM2. Although genomic clones for GSTM3 have not been reported, the cDNA sequences shown in Figure 2 indicate that, among Mu-Class transferase, GSTM4 shares least similarity with GSTM3.

Mapping of hGSTM4

The hGSTM4 gene was mapped to human chromosome 1 p by PCR using the somatic-cell-hybridization methods described by Zhong et al. (1993). In the present study the PCR was carried out using the GSTM4 gene-specific oligo-primer P3(M4) in conjunction with the more general Mu-Class specific oligo-primer P4, which hybridizes to hGSTM1, M2 and M4 genes but not hGSTM3. The expected 2.4 kb PCR product for GSTM4 was detected in certain cells containing human DNA, but not in any of the parental rodent cell lines. The hGSTM4 gene-specific PCR product was from three of the 24 hybrid cell lines (Figure 6). A chromosome assignment of hGSTM4 can be made by correlating the presence of the 2.4 kb PCR product with the human chromosome content of each hybrid cell line. The data obtained from the panel of cells shown in Table 2 strongly suggest an assignment of hGSTM4 to chromosome 1. None of the other chromosomes were found to segregate with the GSTM4-specific PCR product. Most significantly, we found no evidence that the hGSTM4 gene maps to chromosome 3, in contrast with the report of Islam et al. (1989). It is still possible, however, that GSTM3 which exhibits lowest sequence similarity to other Mu-Class GST maps to this chromosome. The finding of Taylor et al. (1990, 1991) that GSTmu2 (GSTM4) and GSTM2 were on the same cosmid clone indicates that the highly similar M1, M2 and M4 genes all map to chromosome 1. The finding that GSTM4 amplified from the hybrid F4Sc13C112, which contains only part of chromosome 1 i.e., 1 p, is consistent with the hGSTM4 gene being located on chromosome 1 p.

Concluding remarks

In the present study, Mu-Class GST from two lymphoblastoid cell lines have been characterized. We have isolated cDNA encoding a novel transferase, present in both cell lines, which has been designated 'GSTM4'. This cDNA has been sequenced along with that of GSTM1b, a polymorphic transferase present in only one of the two cell lines. The gene for GSTM4 has been amplified by PCR and sequenced. Human Mu-Class GST genes are structurally very similar to the gene characterized in the rat. Further analysis of this gene cluster should provide important information on the functions and evolution of Mu-Class GSTs.

REFERENCES

Abramovitz, M. and Listowsky, I. (1987) J. Biol. Chem. **262**, 7770–7773

Board, P. G. (1981) Am. J. Hum. Genet. **33**, 36–43

Board, P. G., Coggan, M., Johnston, P., Ross, V., Suzuki, T. and Webb, G. (1990) Pharmacol. Ther. **48**, 357–369

Bobrow, M. and Cross, J. (1974) Nature (London) **251**, 77–79

Campbell, E., Takahashi, Y., Abramovitz, M., Peretz, M. and Listowsky, I. (1990) J. Biol. Chem. **265**, 9188–9193

Chasseaud, L. F. (1979) Adv. Cancer Res. **29**, 175–274

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979) Biochemistry **18**, 5294–5299

De Jong, J. L., Chang, C-M., Whang-Peng, J., Knutsen, T. and Tu, C-P. D. (1988) Nucleic Acids Res. **16**, 8541–8554

De Klein, A., Van Kessel, A. G., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bottsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J. and Stephenson, J. R. (1982) Nature (London) **300**, 765–767

Edwards, Y. H., Parker, M., Povey, S., West, L. F., Farrington, F. M. and Solomon, E. (1985) Ann. Hum. Genet. **49**, 101–109

Evans, H. J., Buckton, K. E. and Summer, A. T. (1971) Chromosoma **85**, 310–325

Goodfellow, P. N., Banting, G., Levy, R., Povey, S. and McMichael, A. (1980) Somat. Cell Genet. **6**, 777–787

Goodfellow, P. N., Banting, G., Trowsdale, J., Chambers, S. and Solomon, E. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 1190–1194

Harris, H. and Hopkinson, D. A. (1976) Handbook of Enzyme Electrophoresis in Human Genetics, North-Holland Publishing Company, Amsterdam

Hayes, J. D. and Wolf, C. R. (1988) in Glutathione Conjugation – Mechanisms and Biological Significance (Sies, H. and Ketterer, B., eds.), pp. 315–355, Academic Press, London

Hayes, J. D., Strange, R. C. and Percy-Robb, I. W. (1981) Biochem. J. **197**, 491–502

Hayes, J. D., Pickett, C. B. and Mantle, T. J. (eds.) (1990) Glutathione S-transferases and Drug Resistance, Taylor and Francis, London

Heistercamp, N., Groffen, J., Spurr, N. K., Goodfellow, P. N., Solomon, E., Carritt, B. and Bodmer, W. F. (1982) Nature (London) **299**, 747–749

Hsiao, K.-C. (1991) Nucleic Acids Res. **19**, 2787

Hussey, A. J., Stockman, P. K., Beckett, G. J. and Hayes, J. D. (1986) Biochim. Biophys. Acta **874**, 1–12

Hussey, A. J., Kerr, L. A., Cronshaw, A. D., Harrison, D. J. and Hayes, J. D. (1991) Biochem. J. **273**, 323–332

Islam, M. Q., Platz, A., Spirer, J., Szpirer, L., Levan, G. and Mannervik, B. (1989) Hum. Genet. **82**, 338–342

Kamei, K., Oshino, R. and Hara, S. (1990) J. Biochem. (Tokyo) **107**, 111–117

Lai, H.-C., J., Grove, G. and Tu, C.-P. D. (1986) Nucleic Acids. Res. **14**, 6101–6114

Lai, H.-C. J., Quan, B., Grove, G. and Tu, C.-P. D. (1988) J. Biol. Chem. **263**, 11389–11395

Laisney, V., Van Cong, N., Gross, M. S. and Frezal, J. (1984) Hum. Genet. **68**, 221–227

Liu, L.-F. and Tam, M. F. (1991) Biochim. Biophys. Acta **1090**, 343–344

Mannervik, B., Awasthi, Y. C., Board, P. G., Hayes, J. D., Dillio, C., Ketterer, B., Listowsky, I., Morgenstern, R., Muramatsu, M., Pearson, W. R., Pickett, C. B., Sato, K., Widersten, M. and Wolf, C. R. (1992) Biochem. J. **282**, 305–306

Nabholz, M., Miggiano, V. and Bodmer, W. F. (1969) Nature (London) **223**, 358–363

Pearson, W. R., Windle, J. J., Morrow, J. F., Benson, A. M. and Talalay, P. (1983) J. Biol. Chem. **258**, 2052–2062

Pearson, W. R., Reinhart, J., Sisk, S. C., Anderson, K. S. and Adler, P. N. (1988) J. Biol. Chem. **263**, 13324–13332

Philips, J. R., Shepard, E. A., Povey, S., Davis, M. B., Kelsey, G., Monteiro, M., West, L. F. and Cowell, F. (1985) Ann. Hum. Genet. **49**, 267–274

Povey, S., Jeremian, S. J., Barker, R. F., Hopkinson, D. A., Robson, E. B., Cook, P. J. L., Solomon, E., Bobrow, M., Carritt, B. and Buckton, K. E. (1980) Ann. Hum. Genet. **43**, 241–248

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467

Seidegard, J., Pero, R. W., Miller, D. G. and Beattie, E. J. (1986) Carcinogenesis **7**, 751–753

Seidegard, J., Vorachek, W. R., Pero, R. W. and Pearson, W. R. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 7293–7297

Seidegard, J., Pero, R. W., Markowitz, M. M., Roush, G., Miller, G. and Beattie, J. (1990) Carcinogenesis **11**, 33–36

Spurr, N. K., Goodfellow, P. N. and Docherty, A. J. P. (1987) Ann. Hum. Genet. **51**, 189–194

Stockman, P. K., Beckett, G. J. and Hayes, J. D. (1985) Biochem. J. **227**, 457–465

Strange, R. C., Faulder, C. G., Davis, R. A., Hume, R., Brown, J. A. H., Cotton, W. and Hopkinson, D. A. (1984) Ann. Hum. Genet. **48**, 11–20

Strange, R. C., Matharoo, B., Faulder, C. G., Jones, P., Cotton, W., Elder, J. B. and Deakin, M. (1991) Carcinogenesis **12**, 25–28

Suzuki, T., Coggan, M., Shaw, D. C. and Board, P. G. (1987) Ann. Hum. Genet. **51**, 95–106

Taylor, J. B., Oliver, J., Pemble, S. E. and Ketterer, B. (1990) in Glutathione *S*-transferases and Drug Resistance (Hayes, J. D., Pickett, C. B. and Mantle, T. J., eds.), pp. 242–249, Taylor and Francis, London

Taylor, J. B., Oliver, J., Sherrington, R. and Pemble, S. E. (1991) Biochem. J. **274**, 587–593

Townsend, A. J., Goldsmith, M. E., Pickett, C. B. and Cowan, K. H. (1989) J. Biol. Chem. **264**, 21582–21590

van Heyningen, V., Bobrow, L., Bodmer, W. F., Gardiner, S. E., Povey, S. and Hopkinson, D. A. (1975) Ann. Hum. Genet. **38**, 295–303

Vorachek, W. R., Pearson, W. R. and Rule, G. S. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 4443–4447

Received 3 August 1992/21 September 1992; accepted 24 September 1992

Warholm, M., Guthenberg, C., Mannervik, B., Von Bahr, C. and Glaumann, H. (1980) Acta Chem. Scand. **B34**, 607–621

Widersten, M., Pearson, W. R., Engstrom, A. and Mannervik, B. (1991) Biochem. J. **276**, 519–524

Zhong, S., Howie, A. F., Ketterer, B., Taylor, J. B., Hayes, J. D., Beckett, G. J., Wathen, C. G., Wolf, C. R. and Spurr, N. K. (1991) Carcinogenesis **12**, 1533–1537

Zhong, S., Wolf, C. R. and Spurr, N. K. (1993) Hum. Genet., in the press