Vol. 180, No. 1, 1991
October 15, 1991

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 15-22

THE HUMAN H_b (MU) CLASS GLUTATHIONE *S*-TRANSFERASES ARE ENCODED BY A DISPERSED GENE FAMILY

Jeff L. DeJong, T. Mohandas* and Chen-Pei D. Tu[1]

Department of Molecular and Cell Biology
The Pennsylvania State University, University Park, PA 16802

*Division of Medical Genetics, Harbor-UCLA Medical Center
Torrance, CA 90509

Received August 28, 1991

SUMMARY: The human glutathione *S*-transferases are products of a gene superfamily which consists of at least four gene families. The various glutathione *S*-transferase genes are located on different human chromosomes, and new gene(s) are still being added to the gene superfamily. We have characterized a cDNA in pGTH4 encoding human glutathione *S*-transferase subunit 4 (GSTμ) and mapped its gene (or a homologous family member) on chromosome 1 at p31 by *in situ* hybridization. Genomic Southern analysis with the 3′ noncoding region of the cDNA revealed at least four human DNA fragments with highly homologous sequences. Using a panel of DNAs from mouse-human somatic cell hybrids in genomic DNA hybridization we show that the H_b (or B) genes of human glutathione *S*-transferases are on three separate chromosomes: 1, 6, and 13. Therefore, the glutathione *S*-transferase B gene family, which encodes the H_b (mu) class subunits, is a dispersed gene family. The GSTμ(ψ) gene, whose expression is polymorphic in the human population, is probably located on chromosome 13. We propose that the GSTμ(ψ) gene was created by a transposition or recombination event during evolution. The null phenotype may have resulted from a lack of DNA transposition just as much as from the deletion of an inserted gene.

The glutathione *S*-transferases (GSTs; EC2.5.1.18) are dimeric proteins which are important in drug biotransformation, xenobiotics metabolism, and protection against peroxidative damage (for recent reviews, see refs 1,2). The human GST gene superfamily consists of at least four genes and gene families (3,4). The H_a or A genes encode the alpha (H_a) class subunits; the H_b or B genes encode the mu class (H_b) subunits. In addition, there is probably a single functional gene for the GST π and one for the microsomal GST (4-8). The various GST genes are located on different human chromosomes, and new gene(s) are still being added to the gene superfamily (4,9-12). GSTs from the same gene family have different but overlapping substrate specificities (10). The H_b (mu) class subunits, which are encoded by a multigene family, may be important in nitroglycerin metabolism, among other functions (4,13). There are three highly related H_b cDNAs with nearly identical 3′ noncoding

[1]To whom correspondence should be addressed.

0006-291X/91 $1.50

sequences, and a fourth one with ~75% homology in the coding region but very little homology in the 3′ noncoding region (14). One of the three encodes a muscle-specific GST subunit (15). The other two differ by a single amino acid at position 172 with a lys in the GSTμ (4) and an Asn in GSTψ (16). They are expressed in liver and other tissues and considered to be allelic. Their expression is highly polymorphic in humans with a very high frequency (>50%) of the null phenotype among Caucasian, Chinese and Indian populations (1,17,18). Genomic analysis with 5′ end and 3′ end probes of the GSTμ (subunit 4) cDNA (pGTH4) revealed that a complete gene may be missing in the null-phenotype individuals (16,18). This "deficiency" has been correlated to a possible increase in lung cancer incidence among smokers (19). In this communication we analyzed a panel of mouse-human somatic cell hybrids with an H_b (GSTμ) cDNA probe and report that the GST B (H_b) genes are located on at least three different chromosomes. We also discuss the implications of this dispersed gene family on the molecular basis of polymorphic expression of the mu(H_b) class GST subunits.

EXPERIMENTAL PROCEDURES

Human DNAs and GSTμ cDNA probes. The human DNAs were isolated according to a modification of Blin and Stafford (20) from liver samples provided by Dr. D. S. Chen of National Taiwan University Hospital in Taipei, by Dr. C. C. Juan of the Institute of Biomedical Sciences, Academia Sinica in Taipei, and by NDRI in Philadelphia. The cDNA for GSTμ (human subunit 4), pGTH4, has been described previously (4). The 5′ end probe, an 88 bp *Eco*RI-*Hinf*I fragment, contains 24 codons in GSTμ. The 3′ probe is a 386 bp *Hinf*I-*Eco*RI fragment in the 3′ noncoding sequence. The whole probe (W) consists of the complete pGTH4 cDNA insert in a ~1130 bp *Eco*RI fragment (4,18).

Mouse-human somatic cell hybrids. A panel of 13 mouse-human cell hybrids was used for chromosome localization. This panel was constructed by fusion of thymidine kinase-deficient B82 mouse cells and normal human male fibroblasts (IMR91) as described earlier (21). The chromosome content of each line was analyzed by Q-banding on at least 30 metaphases per hybrid clone and recorded in each line as the percent of cells containing each specific chromosome. A chromosome which appears in less than 15% of the metaphases is regarded as absent from a particular hybrid line. In addition, a Chinese hamster-human somatic cell hybrid (CF11-4) retaining 9 pter→9q34 (22) was used to establish whether human chromosome 9 contains this gene, since the other lines lack this human chromosome.

Restriction digestion and DNA transfer. The genomic DNAs (10-15 μg each) from human livers and somatic cell hybrids were digested to completion with *Hind*III, *Pst*I or *Eco*RI and separated by electrophoresis on 0.8% agarose gels in Tris-acetate buffer (20). After staining with ethidium bromide and photography the gels were processed for transfer onto Biotrans™ nylon membranes (ICN Pharmaceuticals, Irvine, CA) according to the manufacturer's procedure.

Radioactive labeling of DNA probes. This is accomplished by using nick translation in the presence of [α-^{32}P]dCTP to a specific activity of 1 to 5x10^8 dpm/μg. Genomic Southern hybridization was carried out according to Sambrook *et al.* (20) in the presence of 50% formamide at 42°C for 24 to 36 h. The Biotrans™ nylon membrane was washed according to Sambrook *et al.* (20) before exposure in the presence of an intensifying screen at -70°C.

RESULTS

There are at least eight different *Hind*III hybridization patterns with the H_b cDNA probe pGTH4 in the human population (23). However, only two different hybridization patterns can be detected with other restriction enzymes (e.g. *Pst*I, *Eco*RI) (Figure 1) (16,23). The DNA hybridization patterns in lanes 5 and 6 of Figure 1 correspond to the null phenotype in liver GSTμ(ψ) expression (23). For *Eco*RI digestions there are four DNA fragments (10 kb, 8 kb, 5.1 kb, and 4.3 kb) that are positive to the 3′ noncoding region of the GSTμ cDNA probe. This is highly unusual because the 3′ noncoding region is often unique to a member of a gene family and there is no *Eco*RI site or intron in the 3′ noncoding region of the cDNAs. Therefore, this type of pattern suggests that there are four closely related genes or pseudogenes in the human genome.

The genomic hybridization analysis using a panel of 13 mouse-human somatic cell hybrids (*Eco*RI digestion) was carried out with both the complete cDNA and the 3′ noncoding region of pGTH4. Results in Figure 2 indicate that there is substantial homology to pGTH4 cDNA in the mouse genomic DNA and that the hybridization patterns are much too complicated for definitive chromosome assignment. On the other hand, the 3′ noncoding region probe revealed from one to four bands in different somatic cell hybrid lines containing different human chromosomes (lanes 2-5). The fact that these four DNA fragments did not segregate together suggested that each of these four genes be on different chromosomes. Discordancy analysis of chromosome segregation is presented in Table 1. The 5.1 kb and 4.3 kb *Eco*RI fragments co-segregated with chromosome 1, consistent with the *in situ* hybridization results using the pGTH4 cDNA probe (4). The 10 kb *Eco*RI fragment signal, however, did not segregate with chromosome 1. Instead it segregated with chromosome 6 perfectly. The 8 kb *Eco*RI fragment, which encodes the polymorphic GSTμ (GSTψ) subunit segregated with a third chromosome, chromosome 13. We are able to map the chromosome location of this particularly interesting gene because the parental human fibroblast (IMR91) is probably derived from a GSTμ(ψ) positive individual as judged from its *Hind*III hybridization pattern (see pattern of Figure 1, lane 2, panel W).

DISCUSSION

The polymorphic expression of a phase II drug metabolizing enzyme in GSTμ(ψ) has many interesting implications in pharmacogenetics and cancer/environmental epidemiology. The presence/absence of a GST gene not only affects GST function in the liver but also affects GST function in other

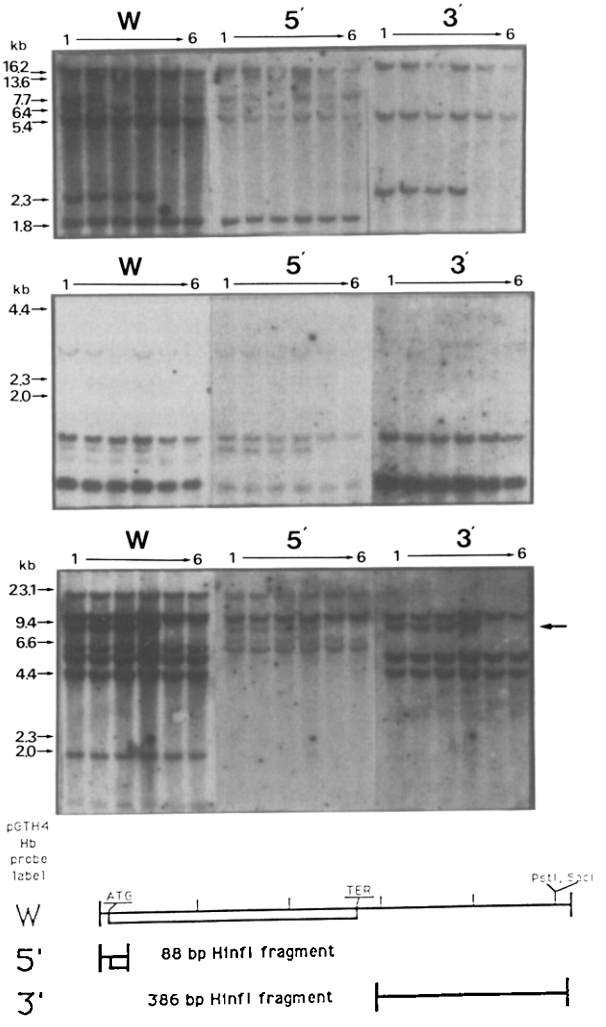

Figure 1. Genomic hybridization analysis of six different human DNAs with portions of GSTμ cDNA probe. W, complete cDNA; 5′, an 88 bp 5′ fragment of the cDNA pGTH4; 3′, a 386 bp fragment of 3′ noncoding region of pGTH4 cDNA. Size markers are in kb units. The same blot was stripped off the previous probe before the next hybridization. The order of probes is W, 5′ and 3′.

tissues or organs (e.g. heart and white blood cells) expressing GSTμ(ψ). The effects should be directly related to the substrate specificity and biochemical functions of GSTμ(ψ) and other regulatory and physiological functions which GSTμ(ψ) may have.

There are four H_b class GST cDNAs characterized by complete sequence analysis: pGTH4 for GSTμ (4), λGTH411 for GSTψ (16), a muscle-specific GST (15), and a brain-specific GST (14). The first three of them have nearly identical 3′ noncoding sequences (4,15,16). If the 8 kb *Eco*RI fragment encodes GSTμ(ψ) and one of the other 3 *Eco*RI fragment in Figure 1 corresponds to the muscle-specific GST H_b (mu), we might still have two B (H_b) genes

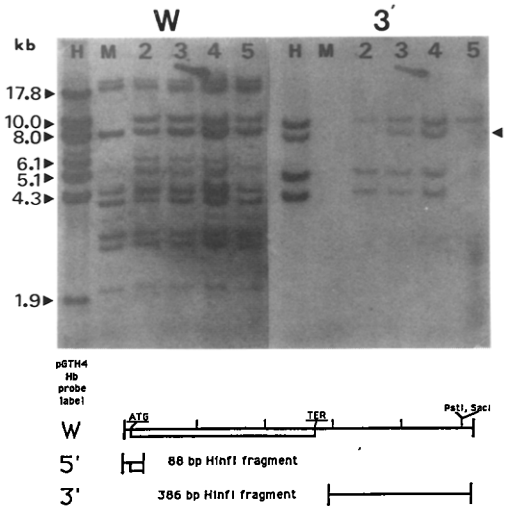

Figure 2. Representative mouse-human somatic cell hybrid genomic DNA hybridization results. Hybridization probe for the left panel (W) was the entire cDNA insert in pGTH4 (*Eco*RI fragment). The probe for the right panel (3′) is the 3′ noncoding sequence of pGTH4 (386 bp). H, human genomic DNA; M, mouse DNA (parental line B-82); lanes 2 to 5 are different hybrid lines 84-2, 84-3, 84-4, 84-5, respectively. Size markers in kb are shown to the left and correspond to the signals from human DNA digested with *Eco*RI. Each lane contains approximately 10 μg of DNA. Hybridization was carried out after transfer to a Biotrans nylon membrane. It was hybridized with probe W first, and followed by probe 3′ after stripping off the previous probe.

(pseudogenes) with highly homologous 3′ noncoding sequence unaccounted for. Either there are two more genes very closely related to GSTμ(Ψ) and the muscle GST or certain gene(s) may have more than one copy on different chromosomes. The 5.1 kb and 4.3 kb *Eco*RI fragments segregate together with chromosome 1. In conjunction with the recent cosmid clone analysis by Taylor *et al.* (24) revealing three linked H_b (mu) genes we suggest that genes represented by these two *Eco*RI fragments (5.1 kb and 4.3 kb) are closely linked. The B (H_b) gene on the 10 kb *Eco*RI fragment of chromosome 6 should correspond to a different member gene or pseudogene, which is not linked to the two on chromosome 1. The gene on chromosome 6 may be the muscle-specific GST gene reported recently (15). The GSTμ(Ψ) gene(s) is most probably located on a third chromosome, chromosome 13, based on the results in Table 1. A majority of the human population apparently do not have this gene and thus are "deficient" for this isozyme expression in the liver and other tissues (17,19,23). Our mapping results differ from the claim by Mannervik *et al.*, who used the rat Y_{b2} cDNA as the homologous probe and mapped the human mu class (B family) gene(s) to chromosome 3 (25). It is possible that they have detected yet another member of the B gene family.

In summary, the H_b (mu) class GST isozymes are encoded by a dispersed multigene family, which would have four highly homologous genes among them.

Table 1. Correlation Between the 5.1 Kb and 4.3 Kb, 10 Kb, and 8 Kb *Eco*RI Hybridization Bands Detected with the GSTμ cDNA Probe and Specific Human Chromosomes

Human Chromosome	Signal/Chromosome Retention														
	5.1 and 4.3 Kb					10 Kb					8 Kb				
	+/+	-/-	+/-	-/+	D(%)[a]	+/+	-/-	+/-	-/+	D(%)[a]	+/+	-/-	+/-	-/+	D(%)[a]
1	6	7	0	0	0	6	1	6	0	46	4	7	0	2	15
2	4	5	2	2	31	5	1	7	0	54	2	6	2	3	39
3	5	4	1	3	31	7	0	5	1	46	4	5	0	4	31
4	5	3	1	4	38	9	1	3	0	23	3	3	1	6	54
5	3	3	3	4	54	7	1	5	0	38	1	3	3	6	69
6	6	1	0	6	46	12	1	0	0	0	4	1	0	8	62
7	5	3	1	4	38	8	0	4	1	38	4	4	0	5	39
8	6	1	0	6	46	11	0	1	1	15	4	0	0	9	69
9	0	7	6	0	46	0	1	12	0	92	0	9	4	0	31
10	4	6	2	1	23	4	1	8	0	62	3	7	1	2	23
11	0	3	6	4	77	4	1	8	0	62	0	5	4	4	62
12	4	3	2	4	46	8	1	4	0	31	2	3	2	6	62
13	4	7	2	0	15	4	1	8	0	62	4	9	0	0	0
14	5	2	1	5	46	9	1	3	0	23	4	4	0	5	39
15	6	3	0	4	31	9	0	3	1	31	4	3	0	6	46
16	2	6	4	1	38	2	1	10	0	77	0	7	4	2	46
17	5	0	1	7	62	11	0	1	1	15	4	1	0	8	62
18	6	5	0	2	15	8	1	4	0	31	4	5	0	4	31
19	6	6	0	1	8	7	1	5	0	38	4	4	0	5	39
20	4	4	2	3	38	7	1	5	0	38	2	4	2	5	54
21	2	3	4	4	62	6	1	6	0	46	1	4	3	5	62
22	4	3	2	4	46	7	1	5	0	38	2	4	2	5	54
X	1	6	5	1	46	1	1	11	0	85	0	7	4	2	46
Y	1	7	5	0	38	2	1	10	0	77	1	8	3	1	31

[a] D = % Discordancy = $\frac{(+/-)+(-/+)}{(+/+)+(-/-)+(+/-)+(-/+)}$

At least two additional H_b (mu) class genes have been isolated in the form of a brain-specific and a testis-specific cDNAs (14,26). Thus, the minimum number of B (H_b) family genes may be six, which is consistent with our genomic Southern analysis using an 88 bp 5′ end probe from the H_b cDNA pGTH4 (18). It is very likely that additional H_b (mu) class genes may be discovered in the near future.

It may not be strictly a deficiency for those individuals missing the GSTμ (ψ) isozyme. That is, in some cases, it may be an advantage not to have the GSTμ (ψ) expression. The origin of GSTμ(ψ) is another intriguing question. Because the majority of human population do not have it and the gene on chromosome 13 is a member of a dispersed gene family, gene deletion for the null phenotype may not be the best explanation. We would like to propose an alternative view. A minor proportion of the human population may have gained a copy of the B (H_b) class GST gene, possibly by DNA transposition or gene conversion/recombination events during evolution. It could be that the target chromosome(s) may vary in different parts of the world, if the chromosome mapping result by Mannervik *et al.* (25) is true. The *Hind*III RFLP may be useful to trace back to the origin(s) of this polymorphism in the human

population. The complete characterization of the GSTμ gene and other B family genes is essential in our understanding of the expression and evolution of the GST gene superfamily, and GST-mediated pharmacogenetics and molecular epidemiology.

ACKNOWLEDGMENTS

We thank Cheryl Meeker, Merry Passage, and Loretta Tu for technical assistance and Eileen McConnell for typing the manuscript. This research project was supported by grants from the National Institute of Environmental Health Sciences of USPHS (ES 02678 and ES 00140) and a Biomedical Research Support Grant to The Pennsylvania State University (2 S07 RR07082).

REFERENCES

1. *Glutathione S-Transferase and Drug Resistance* (Eds. Hayes, J. D., Mantle, T. J., and Pickett, C. B.) Taylor and Francis, London, 1990.
2. Pickett, C. B. and Lu, A. Y. H. (1989) Ann. Rev. Biochem. 58, 743-764.
3. Lai, H.-C. J. and Tu, C.-P. D. (1986) J. Biol. Chem. 261, 13793-13799.
4. DeJong, J. L., Chang, C.-M., Whang-Peng, J., Knutsen, T., and Tu, C.-P. D. (1988) Nucleic Acids Res. 16, 8541-8554.
5. Tu, C.-P. D., Matsushima, A., Li, N., Rhoads, D. M., Srikumar, K., Reddy, A. P., and Reddy, C. C. (1986) J. Biol. Chem. 261, 9540-9545.
6. Rhoads, D. M., Zarlengo, R. P., and Tu, C.-P. D. (1987) Biochem. Biophys. Res. Commun. 145, 474-481.
7. Kano, T., Sakai, M., and Muramatsu, M. (1987) Cancer Res. 47, 5626-5630.
8. DeJong, J. L., Morgenstern, R., Jornvall, H., DePierre, J. W., and Tu, C.-P. D. (1988) J. Biol. Chem. 263, 8430-8436.
9. Board, P. G. and Webb, G. C. (1987) Proc. Natl. Acad. Sci. USA 84, 2377-2381.
10. Chow, N.-W. I., Whang-Peng, J., Kao-Shan, C.-S., Tam, M. F., Lai, H.-C. J., and Tu, C.-P. D. (1988) J. Biol. Chem. 263, 12797-12800.
11. DeJong, J. L., Mohandas, T., and Tu, C.-P. D. (1990) Genomics 6, 379-382.
12. Moscow, J. A., Townsend, A. J., Goldsmith, M. E., Whang-Peng, J., Vickers, P. J., Poisson, R., Legault-Poisson, S., Myers, C. E., and Cowan K. H. (1988) Proc. Natl. Acad. Sci. USA 85, 6518-6522.
13. Tsuchida, S., Maki, T., and Sato, K. (1990) J. Biol. Chem. 265, 7150-7157.
14. Campbell, E., Takahashi, Y., Abramovitz, M., Peretz, M., and Listowsky, I. (1990) J. Biol. Chem. 265, 9188-9193.
15. Vorachek, W. R., Pearson, W. R., and Rule, G. S. (1991) Proc. Natl. Acad. Sci. USA 88, 4443-4447.
16. Seidegard, J., Vorachek, W. R., Pero, R. W., and Pearson, W. R. (1988) Proc. Natl. Acad. Sci. USA 85, 7293-7297.
17. Board, P. G. (1981) Am. J. Hum. Genet. 33, 36-43.
18. DeJong, J. L. and C.-P. D. Tu. in *GSH S-Transferases and Drug Resistance* (Ed. Hayes, J. D., Pickett, C. B., and Mantle, T. J.) Taylor and Francis Ltd., London and Philadelphia, pp. 250-261, 1990).
19. Seidegard, J., Pero, R. W., Miller, D. G., and Beattie, E. J. (1986) Carcinogenesis 7, 751-753.
20. *Molecular Cloning: A Laboratory Manual*, 2nd Edition (Eds. Sambrook, J., Fritsch, E. F., and Maniatis, T.) Cold Spring Harbor Laboratory Press, NY, 1989.

21. Mohandas, T., Heinzmann, C., Sparkes, R. S., Wasmuth, J., Edwards, P., and Lusis, A. J. (1986) Somat. Cell Mol. Genet. 12, 89-94.
22. Mohandas, T., Sparkes, R. S., Sparkes, M. C., and Shulkin, J. D. (1977) Proc. Natl. Acad. Sci. USA 74, 5628-5631.
23. DeJong, J. L. (1990) Ph.D. Dissertation, The Pennsylvania State University, University Park, PA.
24. Taylor, J. B., Oliver, J., Sherrington, R., and Pemble, S. E. (1991) Biochem. J. 274, 587-593.
25. Quamrul Islam, M., Platz, A., Szpirer, C., Levan, G., and Mannervik, B. (1989) Hum. Genet. 82, 338-342.
26. Campbell, E. (1990) Ph.D. Dissertation. Albert Einstein College of Medicine. Bronx. NY.