Hum Genet (1985) 69:337–339

Human Genetics

DNA polymorphism in the 5′ flanking region of the human carbonic anhydrase II gene on chromosome 8

Bailey L. Lee, Patrick J. Venta, and Richard E. Tashian

Department of Human Genetics, Box 015, 4708 MS II, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA

Summary. A restriction-fragment-length polymorphism (RFLP) is described which is associated with the human carbonic anhydrase II gene *(CA2)* that codes for one of the three genetically distinct carbonic anhydrase isozymes, CA I, CA II, and CA III. The isolated DNA was cleaved with several restriction enzymes and subjected to Southern blot hybridization analysis using a DNA probe containing the 5′ end of the human CA II gene. A two allele RFLP which was detected with the restriction endonuclease, *Taq* I, is expressed phenotypically on Southern blots as either a 5.4 kilobase (kb) fragment or as 4.0 and 1.4 kb fragments. These fragments result from the presence or absence of a *Taq* I recognition site in the 5′ flanking region approximately 1.0 kb from the initiation codon of the CA II gene. Segregation analysis showed that the alleles are inherited in a Mendelian fashion, with a frequency of 50%.

Introduction

The multigene carbonic anhydrase (CA) family in humans is made up of three isozymes, designated CA I, CA II, and CA III (Tashian et al. 1983; Hewett-Emmett et al. 1984), and there is evidence for two additional CA genes which code for a membrane-bound (Whitney and Briggle 1982; Wistrand 1984) and a mitochondrial CA isozyme (Dodgson et al. 1980; Storey et al. 1984). At present, the chromosomal localization of the human CA genes is limited to the assignment of the CA II gene, *CA2,* to chromosome 8 (Venta et al. 1983). However, there is some indirect evidence (Carter 1972; DeSimone et al. 1973; Eicher et al. 1976) that the CA I and CA II genes may be linked in humans. A number of variants of the human CA isozymes, detectable at the protein level, are now known, and some of these have been reported to occur as widespread polymorphisms (Tashian et al. 1980, 1983; Hewett-Emmett et al. 1983).

Thirteen genetic markers (some confirmed and some provisional) have been assigned to human chromosome 8 (Meera Khan and Smith 1984). Of these, electrophoretic polymorphisms have been reported for the protein products of only the carbonic anhydrase II and glutathione reductase genes, *CA2* and *GSR* (Venta et al. 1983). Of the four genes on human chromosome 8 for which DNA probes are available [i.e., thyroglobulin *(TG)* (Van Ommen et al. 1984); carbonic anhydrase II *(CA2)* (Venta et al. 1983); and the oncogenes c-*MOS* and c-*MYC* (Neel et al. 1982; Prakash et al. 1982; Dalla-Favera et al. 1982; Sakaguchi et al. 1983)], restriction fragment length polymorphisms (RFLPs) have been reported to be associated with the *TG* and c-*MOS* genes (Van Ommen et al. 1984; D. Barker and R. L. White, cited by Skolnick et al. 1984). In the present report, we describe a DNA polymorphism located about 1 kb (kilobase pair) from the 5′ end of the human CA II gene.

Offprint requests to: Richard E. Tashian, Department of Human Genetics, Box 015, 4708 MS II, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA

Materials and methods

DNA preparation

High molecular weight DNA was prepared by standard techniques (Bell et al. 1981) from 10–20 ml of heparinized fresh blood or outdated blood obtained from the University of Michigan Hospital Blood Bank. A 3 to 5-fold excess of a particular restriction enzyme (Table 1) was used to digest 2–10 μg of DNA under conditions suggested by the supplier for at least 16 h. The resulting fragments were separated in either 0.8% or 2% agarose gels and transferred to nitrocellulose paper by the method of Southern (Southern 1975).

Table 1. Summary of the fragments generated with various restriction enzymes

Enzyme	Fragment pattern	Number of individuals	Number of sites
Bam HI	23.1, 3.2, 0.8	18	4
Bql I	15.8, 7.9	18	3
Pst I	12.2	18	2
Xba I	4.7	18	2
Eco RI	3.8	18	2
Alu I	1.2, 0.4, 0.3	14	4
Hae III	1.4	14	2
Hinf I	2.0, 1.1, 0.7, 0.6, 0.4, 0.2	15	7
Mbo I	1.3, 0.9, 0.8, 0.4	8	5
Msp I	1.6, 0.8, 0.4, 0.3	23	5
Taq I	*5.4, *4.0, 2.1, *1.4, 0.4, 0.3	22	6

Southern transfers of the DNA of unrelated individuals digested with the enzymes indicated were hybridized with the human CA II probe (H25-3.8). The lengths of the restriction fragments observed are in kilobase pairs. The *asterisks* indicate fragments that varied among the individuals. The number of individuals screened with each enzyme is also shown. The estimate of the number of sites examined assumes that all fragments observed were contiguous

Hybridization probe

A recombinant clone (H25) containing a portion of the human CA II gene was isolated from a human genomic lambda phage library (Lawn et al. 1978) provided by Dr. T. Maniatis. The clone was identified by screening the library with a mouse CA II cDNA clone (Curtis et al. 1983) constructed by Dr. P. J. Curtis. A 3.8kb *Eco* RI fragment present in clone H25 was subcloned in pBR325 (Venta et al. 1984) and used as the probe in this study. This fragment contains no repeptitive elements as shown by the presence of a single band on a Southern blot of human genomic DNA cut with *Eco* RI. Partial sequencing of the subclone, designated H25-3.8, indicates that it contains the first two exons of the human CA II gene (Venta et al. 1984; and unpublished results).

Hybridization procedures

Prior to hybridization with the probe, which was nick-translated to $5 \times 10^7 - 2 \times 10^8$ cpm/μg (Maniatis et al. 1982), the filters were prehybridized in 50% formamide, 5 × SSPE (1 × SSPE = 0.15*M* NaCl, 10m*M* NaH_2PO_4, 1m*M* EDTA, pH 7.4), 100μg/ml denatured salmon sperm DNA, 5 × Denhardt's solution (1 × Denhardt's solution = 0.02% bovine serum albumin 0.02% polyvinylpyrrolidone, 0.02% Ficoll), 0.1% SDS (sodium dodecyl sulfate) at 42°C for 6–24h. Hybridization was performed at 42°C for 24–48h with $1–2 \times 10^6$ cpm/ml of ^{32}P-labeled probe in hybridization solution identical to that used for prehybridization except it also contained 10% dextran sulfate. Filters were washed at 50°C first in 2 × SSC (0.3*M* NaCl, 30m*M* sodium citrate) and 0.1% SDS (a 5min wash, followed by a 15min wash), and then a high stringency wash in 0.1 × SSC and 0.1% SDS (a 30min wash, followed by a 15min wash), blotted dry and autoradiographed on Kodak AR5 X-ray film with intensifying screens (DuPont Cronex Lighting Plus) at −70°C for 1–7 days.

Results and discussion

The DNA from several unrelated individuals was analyzed for polymorphisms with the H25-3.8 probe (as described in Materials and methods) after digestion with restriction enzymes that recognize six-base pair sequences (*Eco* RI, *Xba* I, *Pst* I, *Bql*I, and *Bam* HI) as well as four-base pair sequences (*Alu* I, *Hae* III, *Hinf* I, *Mbo* I, *Msp* I, and *Taq* I). All the individual DNA samples screened with a particular enzyme, except *Taq* I, displayed the same banding patterns (Table 1). In the case of *Taq* I, the probe revealed a restriction-site polymorphism with variant bands of 5.4, 4.0, and 1.4kb, and nonpolymorphic bands of 2.1, 0.4, and 0.3kb. Based upon restriction mapping and partial sequencing of H25-3.8, a restriction map of the CA II locus has been constructed (Fig. 1). The *Taq* sites bordering the 0.3 and 0.4kb fragments were defined by restriction mapping and sequencing (unpublished data), and one of the sites was found to be present in exon 2. Because the 4.0kb fragment would not be detected if the polymorphic site was located on the 3′ side of exon 2, the site must be in the 5′ flanking region. The polymorphic site is not present in H25-3.8. An individual homozygous for the absence of the site, −/−, has only the largest (5.4kb) of the three variant bands. However, for an individual who is homozygous for the presence of the site, +/+, the 5.4kb band is absent, and two new bands, 4.0 and 1.4kb, are produced, while for a heterozygote, +/−, all three variant bands are present.

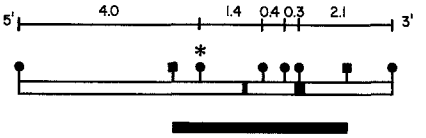

Fig. 1. Restriction map of the 5′ end of the human CA II gene. The locations of the *Taq* I sites (●) and *Eco* RI (■) sites are shown and distances between the *Taq* I sites are in kilobase pairs. The location of the first two exons of the gene are indicated by *solid boxes* and the region covered by the H25-3.8 probe is indicated by the *solid bar*. The polymorphic *Taq* I site (which is not present in H25-3.8) is designated by the *asterisk*

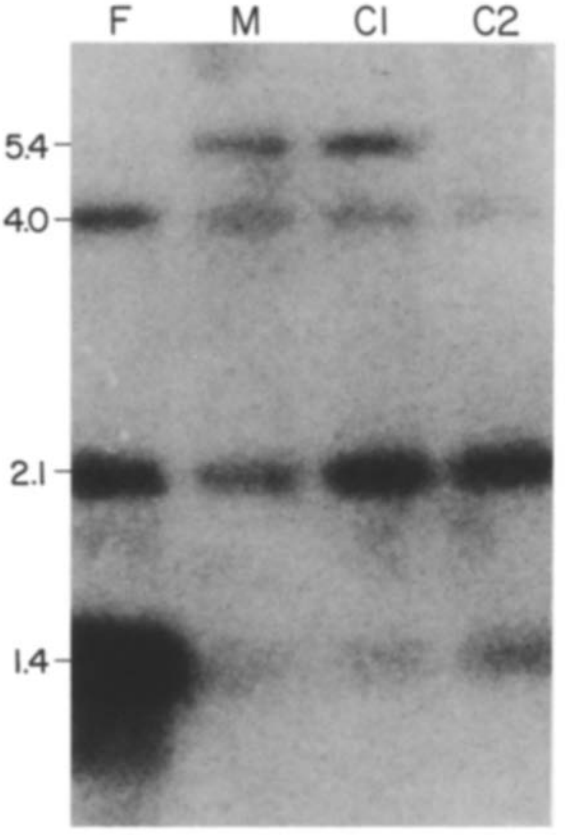

Fig. 2. Inheritance of the *Taq* I polymorphism. The DNAs from four family members were digested with *Taq* I and processed as described under Materials and methods. The autoradiogram shows that the father (F) and one child (C2) are homozygous for the presence of the *Taq* I site while the mother (M) and the other child (C1) are heterozygous. The excessive signal in the 1.4kb band of F is caused by contamination of the sample with pBR322 (data not shown), of which the largest *Taq* I fragments are 1.4kb

The inheritance of this RFLP was ascertained by analyzing the segregation pattern in two white families, one of which (Fig. 2) was informative. The results are consistent with the Mendelian inheritance of the alleles. Twenty-two unrelated individuals were screened for the *Taq* I polymorphism, and the frequency of the allele containing the *Taq* I restriction-site was found to be 50%. The other two RFLPs associated with genes on human chromosome 8, i.e., thyroglobulin *(TG)* and c-*MOS* have frequencies of the new restriction site of 1.5% and 3%, respectively (van Ommen et al. 1984; D. Barker and R. L. White, cited in Skolnick et al. 1984).

The *Taq* 1 RFLP at the CA II locus is due to the presence or absence of a *Taq* I recognition site 1.0kb 5′ from the begin-

ning of the coding region of the gene. That it is not the result of a rearrangement or an insertion or deletion is supported by the fact that no polymorphism was seen with the other 10 enzymes. The enzymes *Taq* I (TCGA) and *Msp* I (CCGG) have been reported to detect RFLPs at a higher rate (Barker et al. 1984) than other enzymes because their recognition sequences contain the dinucleotide CpG. However, in this study no polymorphism was detected with *Msp* I.

The finding of a *Taq* I polymorphism of high frequency associated with the human CA II gene makes genetic screening of informative families with diseases linked to this polymorphism more feasible. In particular, the recently described inherited deficiency of CA II in erythrocytes which manifests itself clinically as a recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification (Sly et al. 1983) can now be subjected to linkage analysis with the *Taq* I polymorphism, and may subsequently lead to a means for the prenatal diagnosis of this deficiency.

Acknowledgements. This work was supported by NIH Grant GM-24681. During this work, Bailey L. Lee was the recipient of a University of Michigan Student Research Fellowship. We thank Dr. David Hewett-Emmett for his helpful advice during this study. We are also grateful to Dr. Peter J. Curtis for supplying us with the mouse CA II cDNA, Dr. Tom Maniatis for the human lambda library, and Dr. James M. Wilson for obtaining the blood samples for the family study.

References

Barker D, Schafer M, White R (1984) Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138

Bell GI, Karam JH, Rutter WJ (1981) Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. Proc Natl Acad Sci USA 78:5759–5763

Camerino G, Grzeschik KH, Jaye M, DeLaSalle H, Tolstoshev P, Lecocq JP, Heilig R, Mandel JL (1984) Regional localization on the human X chromosome and polymorphism of the coagulation factor IX gene (Hemophilia B locus). Proc Natl Acad Sci USA 81:498–502

Carter ND (1972) Carbonic anhydrase isozymes in *Cavia Purcellus, Cavia aperea* and their hybrids. Comp Biochem Physiol B 43:743–747

Curtis PJ, Withers E, Demuth D, Watt R, Venta PJ, Tashian RE (1983) The nucleotide sequence and derived amino acid sequence of cDNA coding for mouse carbonic anhydrase II. Gene 25:325–332

Dalla-Favera R, Bregni M, Erickson J, Patterson D, Gallo RC, Croce CM (1982) Human c-*myc onc* gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci USA 79:7824–7827

DeSimone J, Linde M, Tashian RE (1973) Evidence for linkage of carbonic anhydrase isozyme genes in the pig-tailed macaque, *Macaca nemestrina.* Nature 242:55–56

Dodgson SJ, Forster RE, Storey BT, Mela L (1980) Mitochondrial carbonic anhydrase. Proc Natl Acad Sci USA 77:5562–5566

Eicher EM, Stern RH, Womack JE, Davisson MT, Roderick TH, Reynolds SC (1976) Evolution of mammalian carbonic anhydrase loci by tandem duplication: close linkage of *Car-1* and *Car-2* to the centromere region of chromosome 3 of the mouse. Biochem Genet 14:651–660

Hewett-Emmett D, Welty RJ, Tashian RE (1983) A widespread silent polymorphism of human carbonic anhydrase III (31 Ile↔Val): implications for evolutionary genetics. Genetics 105:409–420

Hewett-Emmett D, Hopkins PJ, Tashian RE, Czelusniak J (1984) Origins and molecular evolution of the carbonic anhydrase isozymes. Ann NY Acad Sci 429:338–358

Lawn RM, Fritsch EF, Parker RC, Blake G, Maniatis T (1978) The isolation and characterization of linked δ- and β-globin genes from a cloned library of human DNA. Cell 15:1157–1174

Meera Kahn P, Smith M (1984) Report of the committee on the genetic constitution of chromosomes 7, 8 and 9. Cytogenet Cell Genet 37:71–102

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory

Neel BG, Jhanwar SC, Chaganti RSK, Hayward WS (1982) Two human c-*onc* genes are located on the long arm of chromosome 8. Proc Natl Acad Sci USA 79:7842–7846

Prakash K, McBride OW, Swan DC, Devore SG, Tronick SR, Aaronson SA (1982) Molecular cloning and chromosomal mapping of a human locus related to the transforming gene of Moloney murine sarcoma virus. Proc Natl Acad Sci USA 79:5210–5214

Sakaguchi AY, Lalley PA, Naylor SL (1983) Human and mouse cellular *myc* protooncogenes reside on chromosomes involved in numerical and structural aberrations in cancer. Somatic Cell Genet 9:391–405

Skolnick MH, Willard HF, Menlove LA (1984) Report of the committee on human gene mapping by recombinant DNA techniques. Cytogenet Cell Genet 37:210–273

Sly WS, Hewett-Emmett D, Whyte MP, Yu Y-SL, Tashian RE (1983) Carbonic anhydrase II deficiency identified as the primary defect in the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. Proc Natl Acad Sci USA 80:2752–2756

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Storey BT, Dodgson SJ, Forster RE (1984) Mitochondrial carbonic anhydrase: the purified enzyme. Ann NY Acad Sci 429:210–211

Tashian RE, Kendall A, Carter ND (1980) Inherited variants of human red cell carbonic anhydrases. Hemoglobin 4:635–651

Tashian RE, Hewett-Emmett D, Goodman M (1983) On the evolution and genetics of carbonic anhydrases I, II and III. In: Ratazzi M, Scandalios JG, Whitt GS (eds) Isozymes, vol 7, Current Topics in Biological and Medical Research. Alan R Liss, New York, pp 79–100

van Ommen GBJ, Bass F, Arnberg AC, Pearson PL, DeVijlder JJM (1984) Chromosome mapping and polymorphism study of the human thyroglobulin gene. Cytogenet Cell Genet 37:562

Venta PJ, Shows TB, Curtis PJ, Tashian RE (1983) Polymorphic gene for human carbonic anhydrase II: a molecular disease marker located on chromosome 8. Proc Natl Acad Sci USA 80:4437–4440

Venta PJ, Montgomery JC, Wiebauer K, Hewett-Emmett D, Tashian RE (1984) Organization of the mouse and human carbonic anhydrase II genes. Ann NY Acad Sci 429:309–323

Whitney PL, Briggle TV (1982) Membrane-associated carbonic anhydrase purified from bovine lung. J Biol Chem 257:12056–12059

Wistrand PJ (1984) Properties of membrane-bound carbonic anhydrase. Ann NY Acad Sci 429:195–206

Received August 31, 1984 / Revised December 17, 1984