FEBS Letters 351 (1994) 411–415

FEBS 14508

Molecular characterization of a class IV human alcohol dehydrogenase gene (*ADH7*)

Shozo Yokoyama*, Yoshinori Matsuo, Rebecca Ramsbotham, Ruth Yokoyama

Department of Biology, Syracuse University, 130 College Place, Syracuse, NY 13244, USA

Received 4 August 1994

Abstract Class IV alcohol dehydrogenase (ADH) is a form preferentially expressed in stomach. We report here the isolation and sequence determination of a novel human ADH gene (*ADH7*). Phylogenetic analysis strongly suggests that *ADH7* is a functional class IV ADH gene.

Key words: Alcohol dehydrogenase; DNA; Genomic; Molecular evolution; Human

1. Introduction

Human alcohol dehydrogenases (ADHs) are dimeric metalloenzymes and had been grouped early into three classes, I–III, according to different electrophoretic properties and different substrate specificities (e.g., see [1]). These ADHs have been conserved among diverse species such as mammals, plants and yeasts [2,3]. Additional ADH classes IV [4] and V [5] have since then discovered. Phylogenetic analysis of available amino acid sequences has shown that class I and V are most closely related, their common ancestor diverged from the ancestor of class II, and their common ancestor diverged from the ancestor of class III before that ([6]; see also [7]).

At present, mammalian class IV and V ADHs are relatively poorly understood. The class V protein has been deduced from genomic and cDNA clones [5], but the protein has not been isolated from natural sources. The class IV protein appears to be expressed mainly in the stomach mucosa [4, 8]. The molecular structure of a class IV protein from rat has been characterized only recently [8], but that for human is not known. We report here the cloning and sequencing of a novel human ADH gene (*ADH7*). Phylogenetic analysis of the amino acid sequence deduced from *ADH7* and other vertebrate ADHs strongly suggests that *ADH7* is a functional class IV ADH gene.

2. Materials and methods

2.1. λ clone characterization and DNA sequencing

Two sets of human genomic DNA libraries were constructed using genomic DNA partially digested with *Sau*3A or *Eco*RI and by ligating it with λEMBL3 or λEMBL4 DNA, respectively (e.g. see [9]). To characterize a novel ADH gene, *ADH7*, two overlapping clones λ1.1 and λ18 have been obtained. λ18 was isolated from the λEMBL4 library using bADH cDNA probe [10]. λ1.1 was isolated from the λEMBL3 library using a 0.5 kb fragment encompassing exons 3 and 4 of λ18 as a probe. λ1.1 contained exons 2–8, while λ18 contained exons 1–5 (Fig. 1). The two clones were verified to be overlapping by restriction enzyme analysis and sequencing. By amplifying genomic DNA using a primer 100 bp from the end of λ1.1 and a degenerate class I exon 9 primer, it became clear that exon 9 is located approximately 1.2 kb from the end of λ1.1. Furthermore, genomic Southern hybridization indicated that exon 9 was contained on a 2.4 kb *Hin*dIII fragment (data not shown). The sequence of exon 9 was obtained by using the inverse PCR method [11]. To perform inverse PCR, 2 oligonucleotides (int8A and int8B) were designed from intron 8 sequence to flank and extend into the unknown genomic region upon circularization of the *Hin*dIII-digested genomic DNA (Fig. 1). Int8A (5′-GGCCCAATGAGTGAAAGTCATTTC-3′) corresponds to sequence 100 bp from the end of λ1.1 and was oriented to amplify towards exon 8. Int8B (5′-CTCAGATACCCTTATCAGAGTTGCG-3′) corresponds to sequence about 200 bp upstream from exon 9 and was oriented to amplify towards exon 9. Ligation of 1 μg *Hin*dIII-digested genomic DNA at a concentration of 2 μg/ml was done to promote monomeric circularization. The ligated product was amplified using int8A and int8B for 30 cycles (94°C for 45 s, 55°C for 90 s, 72°C for 2 min). The expected 1.2 kb product was gel isolated, blunt-ended by Klenow treatment and subcloned into *Eco*RV-digested Bluescript. Exon 9 sequence was obtained using int8B primer. Subcloning and double-stranded sequencing of the λ1.1 and λ18 were conducted as previously described [9,12,13].

2.2. Data analysis

The amino acid sequence, Hsa7, deduced from *ADH7* was compared to other vertebrate ADHs. Vertebrate ADHs considered were class I ADHα (HsaA), β (HsaB), γ (HsaC) from human and those from baboon (PhaB), rhesus monkey (MacA), mouse (MmuA), rat (RnoA), quail (Cja), chicken (Gga), and frog (Rpe), and ADH-E (EcaE) and ADH-S (EcaS) from horse, one class II ADH from human (HsaP), class III ADHs from human (HsaX), mouse (MmuX), rat (RnoX) and horse (EcaX), and class V ADH (Hsa6) deduced from human *ADH6* ([6] and references therein). Furthermore, ADH1 (PmaA) and ADH2 (PmaIV) from deer mouse (*Peromyscus maniculatus*) [14] and class IV ADH from rat (RnoIV) [8] were also included.

To construct a rooted phylogenetic tree for these ADHs, ADH1 and ADH2 from maize (see [6] and references therein) were used as an outgroup. These amino acid sequences can be aligned easily (see [6]). The number (K) of amino acid substitutions per site for two polypeptides was estimated by $K = -\ln(1-p-p^2/5)$, where p is the proportion of different amino acids between the two sequences [15]. Topology and branch lengths of the phylogenetic tree were evaluated by applying the neighbor-joining (NJ) method to the K values and reliability of its topology was evaluated by the bootstrap analysis, as previously described [6].

3. Results and discussion

3.1. Nucleotide sequence of ADH7

ADH7 has 9 exons and 8 introns (Fig. 2), as in the case of *ADH1* [13], *ADH2* [12, 16], *ADH3* [17], *ADH4* [18], and *ADH5* [19], all of which have identical intron positions. Compared to these genes, *ADH6* lacks exon 9, but positions of its introns 1–7 are identical to those of the other ADH genes [5]. The nine exons of *ADH7* span about 22 kb and the approximate sizes of the introns 1–8 of *ADH7* are 5.3, 0.6, 0.3, 0.1, 7.2, 1.4, 3.4,

*Corresponding author. Fax: (1) (315) 443 2012. INTERNET: syokoyam@suvm.syr.edu BITNET: SYOKOYAM@SUVM.BITNET

SSDI 0014-5793(94)00895-7

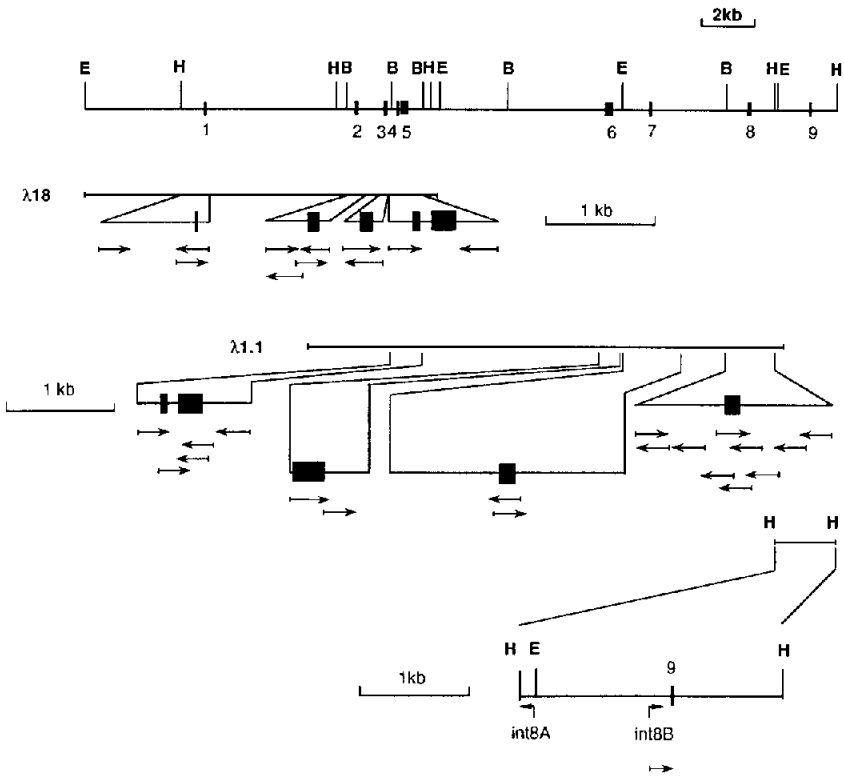

Fig. 1. The genomic composite of the human *ADH7* gene determined by overlapping λ clones and PCR amplification. The nine coding exons are indicated by solid boxes with an enlarged scale being provided under each λ clone and genomic fragment used for inverse PCR amplification. (B) *Bam*HI; (E) *Eco*RI; and (H) *Hin*dIII.

and 2.1 kb, respectively. The GT/AG consensus intron splice junctions are conserved.

The amino acid sequence, Hsa7, deduced from *ADH7* consists of 374 residues (Fig. 2). This size is similar to other ADHs, i.e. 374 residues of HsaX, 375 residues of HsaA, B, and C, 380 residues of HsaP, and 368 residues of Hsa6. Hsa7 shows that functionally important residues are conserved. They include amino acid residues Cys (47), His (68), and Cys (174), which provide ligands to one catalytic zinc atom, and amino acid residues Cys (98), Cys (101), Cys (104), and Cys (112), which provide ligands to another catalytic zinc atom [20]. The sequence conservation of *ADH7* is unlike the *ADH5* pseudogene,

Table 1
Class IV-specific amino acid polymorphism

	71	80	166	196	233	251	261	263	287	317	322	343
EcaX	A	G	L	A	K	Q	G	D	G	T	W	D
HsaX	–	–	–	–	–	–	–	–	–	–	–	–
RnoX	–	–	–	–	–	–	–	–	–	–		–
MmuX	–	–	–	–	–	–	–	–	–	–	–	–
HsaP	–	–	–	–	–	–	–	–	–	–		–
PmaV	T				–	–	–	–	–	S	–	–
Hsa6	–	–	–	–	Q	–	–	–	–	S	–	–
Hsa7	T	–	P	–	M	S	N	G	V	C	L	–
RnoIV	V	E	P	V	L	S	T	Q	–	C	–	G
Rpe	–	–	–	–	–	H	–	–	–	A		–
Gga	–	K	–	–	–	H	–	–	–	S	–	–
Cja	–	K	–	–	–	H	–	–	–	S	–	–
PmaA						–	–	–	–	A	F	E
RnoA	–	–	–	–	–	–	–	–	–	A	F	E
MmuA		–								A	F	–
EcaS	–	–	–	–	–	–	–	–	–	A	F	–
EcaE	–	–	–	–	–	–	–	–	–	A	F	
HsaC	–	–	–	–	–	–	–	–	–	A	F	–
PhaB	–	–	–	–	–	–	–	–	–	A	F	–
HsaB										A	F	
HsaA	–	–	–	–	–	–	–	–	–	A	F	–
MmuA	–	–	M	–	–	–	–	–	–	A	F	–

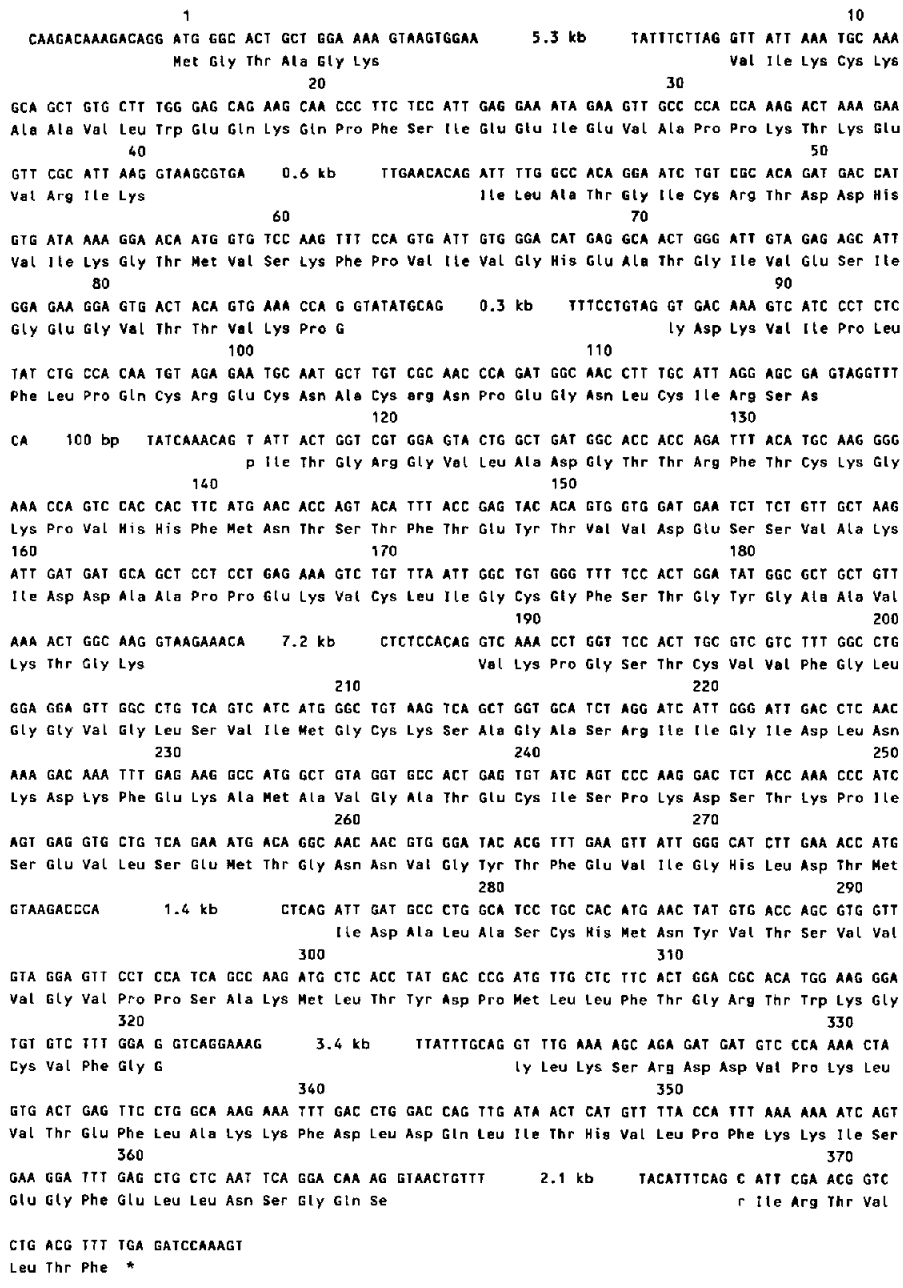

Fig. 2. The nucleotide sequence of *ADH7* and the deduced amino acid sequence.

which contains no introns and three frameshift mutations in the coding region with a total of 10 internal termination codons [9]. The gene duplication event between ADH5 and the pseudogene seems to have occurred only about 3.5 million years ago and the pseudogene has undergone a rapid change since then. Thus, possessing no apparent structural defect and conservation of functionally important codons, *ADH7* seems capable of being functional.

3.2. *A phylogenetic tree*

The phylogenetic tree for the Hsa7 and other vertebrate ADHs is shown in Fig. 3. Fig. 3 clearly shows that Hsa7 is most closely related to RnoIV and belongs to class IV. Bootstrap support of the two ADHs forming one cluster is 1.0 and is highly reliable. Fig. 3 shows that class IV is most closely related to class I, their common ancestor diverged from the ancestor of class V, their common ancestor diverged from the ancestor of class II, and their common ancestor diverged from the ancestor of class III, which is consistent with Fig. 1 in Yokoyama and Harry [6]. Bootstrap analysis shows that clusterings of the mammalian classes I, III, IV and V ADHs are highly reliable. However, the relationship between the classes I and IV in Fig. 3 is not clear-cut and frog, chicken, and quail ADHs may be more closely related to the mammalian class I ADHs than to the class IV ADHs, reflecting the organismal relationship of the three species. Fig. 3 also clearly shows that recently characterized ADH1 (PmaA) and ADH2 (PmaV) from deer mouse belong to classes I and V, respectively.

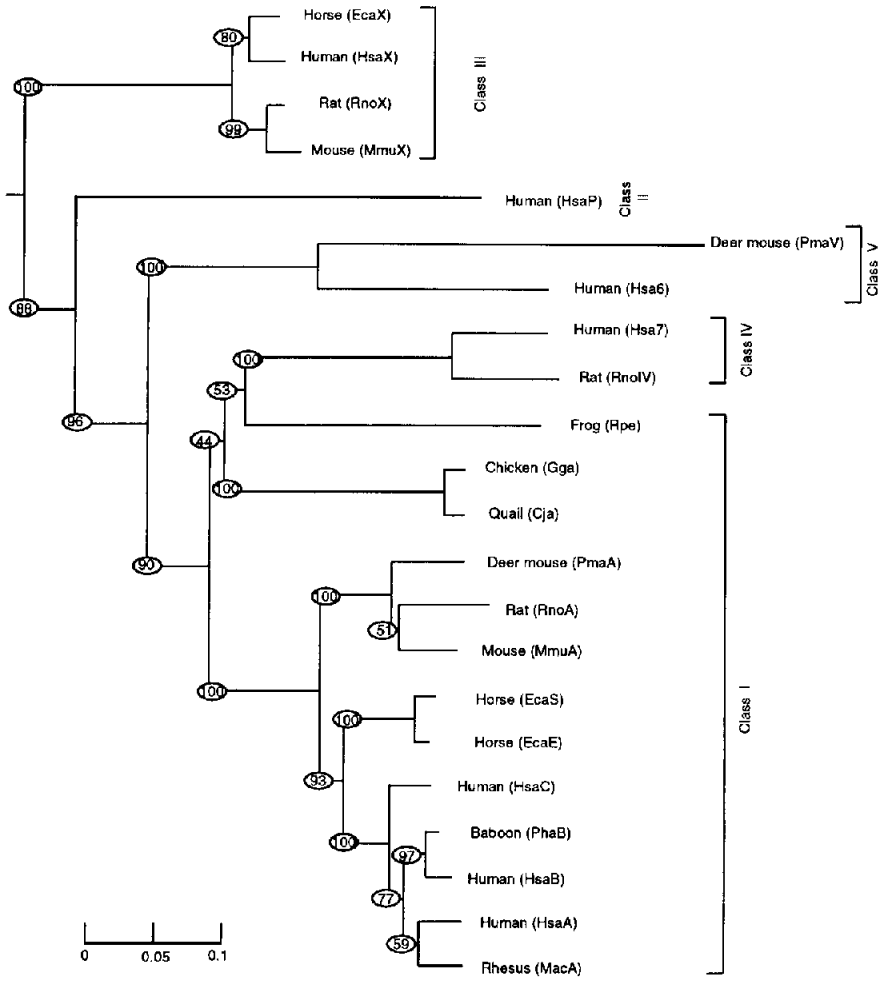

Fig. 3. The rooted phylogenetic tree for the vertebrate ADHs. Branch length was estimated by the number of amino acid replacements per site.

From Fig. 3, we can see that Hsa7 and RnoIV have similar branch lengths after their divergence from each other, showing that the two proteins are evolving with similar rates. The branch length for Hsa7 since its divergence from RnoIV is 0.07 per site. After their divergences from the corresponding murine ADHs, the branch lengths for HsaA, B, and C are 0.10, 0.09, and 0.08 per site, respectively, while that for HsaX is 0.03. Thus, class IV ADHs are evolving with rates somewhere between those of class I and class V ADHs (see also Fig. 3) Interestingly, the zinc-containing ADH from *Drosophila melanogaster* [21] has been found recently and is shown to belong to the evolutionarily most conserved class III. All of these observations strongly suggest that Hsa7 has had a strong functional constraint against amino acid changes and that *ADH7* is a functional gene.

3.3. Characteristics of class IV proteins

By comparing RnoIV to other classes of ADHs, amino acid residues Gly (47), Ala (294), and Tyr (363) of RnoIV are suggested to be class IV-specific changes by affecting both substrate and coenzyme binding [8]. However, the corresponding amino acids for Hsa7 are Arg (48), Val (294), and Asn (363) and the three residues are not shared by the two class IV ADHs.

Having two class IV proteins, RnoIV and Hsa7, it is of interest to identify class IV-specific amino acid changes. Table 1 shows polymorphic residues, where (i) amino acid changes of at least one of the two class IV proteins are not shared by others and (ii) amino acids among other classes of ADHs are either identical or belong to the same functional group. Comparing all vertebrate ADHs, Hsa7 and RnoIV are uniquely distinguished from others by Pro (166), Ser (251), and Cys (317) (Table 1). These three and additional amino acid changes in Hsa7 and RnoIV in Table 1 might have been important in its functional differentiation from class I ADH.

Acknowledgements: This work was supported by the National Institute of Health Grant AA-03539.

References

[1] Smith, M. (1986) Adv. Hum. Genet. 14, 249–290.
[2] Eklund, H., Branden, C.-I. and Jornvall, H. (1976) J. Mol. Biol. 102, 61–73.
[3] Yokoyama, S., Yokoyama, R., Kinlaw, C.S. and Harry, D.E. (1990) Mol. Biol. Evol. 7, 143–154.
[4] Moreno, A. and Pares, X. (1991) J. Biol. Chem. 266, 1128–1133.
[5] Yasunami, M., Chen, C.-S. and Yoshida, A. (1991) Proc. Natl. Acad. Sci. USA 88, 7610–7614.
[6] Yokoyama, S. and Harry, D.E. (1993) Mol. Biol. Evol. 10, 1215–1226.

[7] Sun, H.-W. and Plapp, B. (1992) J. Mol. Evol. 34, 522–535.
[8] Pares, X., Cederlund, E., Moreno, A., Hjelmqvist, L., Farres, J. and Jornvall, H. (1994) Proc. Natl. Acad. Sci. USA 91, 1893–1897.
[9] Matsuo, Y. and Yokoyama, S. (1990) Am. J. Hum. Genet. 46, 85–91.
[10] Yokoyama, S., Yokoyama, R. and Rotwein, P. (1987) Jpn. J. Genet. 62, 241–256.
[11] Ochman, H., Medhora, M.M., Garza, D. and Hartl, D.L. (1990) in: PCR protocols: a guide to methods and applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J., Eds.) pp. 219–227, Academic Press, New York.
[12] Matsuo, Y., Yokoyama, R. and Yokoyama, S. (1989) Eur. J. Biochem. 183, 317–320.
[13] Matsuo, Y. and Yokoyama, S. (1989) FEBS Lett. 243, 57–60.
[14] Zheng, Y.-W., Bey, M., Liu, H. and Felder, M.R. (1993) J. Biol. Chem. 268, 24933–24939.
[15] Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.
[16] Duester, G., Smith, M., Bilanchone, V. and Hatfield, G. W. (1986) J. Biol. Chem. 261, 2027–2033.
[17] Yokoyama, S., Matsuo, Y., Rajasekharan, S. and Yokoyama, R. (1992) Jpn. J. Genet. 67, 167–171.
[18] von Bahr-Lindstrom, H., Jornvall, H. and Hoog, J.-O. (1991) Gene 103, 269–274.
[19] Hur, M.-W. and Edenberg, H.J. (1992) Gene 121, 305–311.
[20] Eklund, H., Nordstrom, B., Zeppezauer, E., Soderberg, B.-O., Tapia, O., Branden, C.-I. and Akeson, A. (1976) J. Mol. Biol. 102, 27–59.
[21] Danielsson, O., Atrian, S., Luque, T., Hjelmqvist, L., Gonzalez-Duarte, R. and Jornvall, H. (1994) Proc. Natl. Acad. Sci. USA 91, 4980–4984.