Molecular characterization of cDNA clones encoding the human alcohol dehydrogenase β_1 and the evolutionary relationship to the other Class I subunits α and γ

BY Shozo YOKOYAMA*, Ruth YOKOYAMA* and Peter ROTWEIN**

**Department of Ecology, Ethology, and Evolution, University of Illinois, Shelford Vivarium, 606 East Healey Street, Champaign, Illinois 61820*

***Department of Medicine, Washington University School of Medicine, St. Louis, Missouri 63110*

(Received May 20, 1987)

ABSTRACT

Three different sizes of cDNA clones coding for the β_1 subunit of human alcohol dehydrogenase (ADH) were isolated from a human cDNA library constructed in bacteriophage λ gt11 and completely sequenced. Comparisons of these three with two other published ADH β_1 cDNA sequences show one silent nucleotide substitution at the third position of codon 220 and two nucleotide differences within the 1337 nucleotides of the 3' untranslated region. Hybridization of a 3' untranslated region of one of the β_1 cDNA clones to DNAs from various individuals has shown a Rsa I restriction fragment length polymorphism. A phylogenetic tree for class I human ADHs α , β , and γ shows that the subunits α and β diverged most recently and that their common ancestor diverged from the ancestor of the subunit γ earlier. This tree topology differs from that based on equal rates of nucleotide (or amino acid) substitution of the three ADH subunits, in which the subunits β and γ are considered to be most recently diverged. The evolutionary rates of nucleotide substitution for the three subunits reveal that the subunit γ is evolving with the slowest rate, followed by β and α , in that order, implying that the subunit γ may be providing the original function of ethanol metabolism.

1. INTRODUCTION

Alcohol dehydrogenases (ADHs) are enzymes that oxidize a variety of primary, secondary, and tertiary aliphatic alcohols and a limited number of cyclic alcohols, in the presence of NAD, to generate the corresponding aldehydes (Smith *et al.* 1986). In the human liver, class I ADH exhibits high activity for ethanol metabolism. The class I ADH contains several homo- and heterodimers, formed by three types of subunits α , β , γ , which are coded by the three separate structural loci, ADH1, ADH2, and ADH3, respectively (Smith *et al.* 1972, 1973a, b).

The majority of Orientals have the "atypical" ADH2² allele, producing the β_2 subunit, while most Caucasians have the "typical" β_1 subunit encoded by the ADH2¹ allele (Stamatoyannopoulos *et al.* 1975). The ADH3 locus is also

polymorphic, although the frequencies of the ADH3¹ allele (codes for γ_1 subunit) and the ADH3² allele (for γ_2 subunit) are almost equally common in Caucasian (Smith *et al.* 1973a, b) and Oriental populations (Watanabe *et al.* 1975; Teng *et al.* 1979; Harada *et al.* 1980). The ADH1 locus is considered to be monomorphic (Smith *et al.* 1972, 1973a, b).

The amino acid sequences of β and γ subunits of ADH have been determined after purification of enzyme activity from human liver (Hempel *et al.* 1984; Buhler *et al.* 1984). More recently, full length cDNA clones for the α (Ikuta *et al.* 1986; Von Bahr-Lindstrom *et al.* 1986), β (Von Bahr-Lindstrom *et al.* 1986; Ikuta *et al.* 1986; Heden *et al.* 1986), and γ (Ikuta *et al.* 1986) subunits have been characterized and the complete amino acid sequences of all three subunits have been deduced.

In this paper, we report DNA sequences of three cDNA clones which also code for the β_1 subunit of human ADH and evaluate the level of nucleotide diversity of the β_1 subunit by comparing currently available cDNA clones. We describe a Rsa I restriction fragment length polymorphism found by hybridization of a 3' region of one of the ADH β cDNA clones to genomic DNA from 27 individuals. We also discuss the molecular evolution and phylogeny of the three class I subunits and provide new insight into the evolution of these enzyme subunits.

2. MATERIALS AND METHODS

Materials

Enzymes including restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, DNA polymerases, ribonuclease A, and proteinase K were purchased from commercial suppliers (New England Biolabs and Bethesda Research Laboratories). Nitrocellulose was obtained from Schleicher and Schuell and Millipore. Radionuclides were purchased from New England Nuclear and Amersham; deoxynucleotide triphosphates and dideoxynucleotide triphosphates were purchased from Pharmacia P-L.

Oligonucleotide synthesis

A 21 base oligonucleotide encoding the portion of the human β subunit between amino acids 286 and 292, i.e., 5' TAC GAT GAC GCT TGT GCC ACA 3' (see Duester *et al.* 1984), was produced in the Department of Biological Chemistry, Washington University, St. Louis, on an Applied Biosystems solid phase DNA synthesizer. The probe was 5'-end-labelled using [γ -³²P] ATP and T4 polynucleotide kinase (Maniatis *et al.* 1982).

cDNA cloning

A human liver cDNA library in λ gt11 (kindly provided by S.L.C. Woo and

T. Chandra) was plated on *Escherichia coli* K-12 strain RY1088 (Young and Davis 1983). Duplicate nitrocellulose filters were prepared (Woo 1979) and hybridized at 55°C in buffer containing 6XSSC (1XSSC=150 mM NaCl/15 mM Na citrate, pH 7), denatured salmon sperm DNA (100 µg/ml), 10X Denhardt's solution. Following hybridization, the filters were washed for 15 minutes at 22°C and for 5 min at 55°C in 6XSSC. Positive plaques were rescreened at lower density, purified to homogeneity and DNA was isolated (Helms *et al.* 1985).

DNA sequence analyses

DNA sequencing was performed by the dideoxy-chain-terminating method (Sanger *et al.* 1977; Biggin *et al.* 1983) after subcloning restriction fragments into M13 mp18 and mp19 bacteriophage (Norlander *et al.* 1983). Data analysis was simplified by computer programs run on an Apple II microcomputer (Larson and Messing 1983). Nucleotide alignments were achieved by using the NUCALN program (Wilbur and Lipman 1983).

DNA isolation, gel electrophoresis and hybridization

High molecular weight human genomic DNA was prepared from nuclei of peripheral blood cells following the method of Kan and Dozy (1978). Agarose gel electrophoresis and blotting of the DNA onto nitrocellulose filters followed standard protocols (Southern 1975). Prehybridization and hybridization were carried out at 65°C for 16–22 hours each in 4XSSC, 0.1 M Tris HCl, pH 7.4, 0.5% sodium dodecyl sulfate (SDS) and 5X Denhardt solution. The hybridized filters were washed extensively at room temperature with 0.2XSSC, 0.1% SDS and then one hour with three changes of 0.1XSSC, 0.1% SDS at 50°C. [³²P]-labelled single-stranded probe DNA was prepared from the p1.1–1 subclone (see Figure 1) by following the protocol of Jeffreys *et al.* (1985).

3. RESULTS

Cloning of cDNAs for β Subunits

Approximately 9×10^4 plaques were screened by hybridization with an oligonucleotide probe for the β subunit as described in MATERIALS AND METHODS. Twelve positives, ranging from 1.0 to 1.8 kilobases (kb), were isolated. Among these positives, the four largest clones λADH5 (about 1.35 kb), λADH6 (about 1.7 kb), λADH8 (about 1.75 kb), and λADH11 (about 1.8 kb) were initially sequenced partially from the 3' end. The results revealed three different polyadenylation sites with λADH6 and λADH11 having identical sites. Three clones (λADH5, λADH8, and λADH11) containing different polyadenylation sites were then selected for complete DNA sequence analyses.

Fig. 1. Restriction map of the cDNA clones corresponding to the β_1 subunit of the human liver alcohol dehydrogenase. Only restriction sites pertinent for sequencing are shown. Arrows indicate direction and extent of nucleotide sequence determined by the dideoxy-chain-terminating method.

Nucleotide Sequence Analysis

Fig. 1 depicts the approach to DNA sequencing of the three cDNA and the pertinent restriction enzyme sites of ADH β cDNA. The DNA sequences and amino acid translation of λ ADH5 (λ 5), λ ADH8 (λ 8), and λ ADH11 (λ 11) are shown in Fig. 2, where pUC ADH14 (p14) in Ikuta *et al.* (1986) and a composite sequence (c) of three overlapping clones obtained by Heden *et al.* (1986) are also shown. It is important to note that pUC ADH14 (Ikuta *et al.* 1986) was isolated from the same cDNA library as ours, but that Heden *et al.* (1986) screened a cDNA library from a different individual.

The clones λ ADH5 and λ ADH8 contain short 5' untranslated regions of 19 and 20 nucleotides respectively, the full coding sequence, plus 213 and 591 nucleotides, respectively, of the 3' untranslated region; λ ADH11 consists of 1,803 nucleotides which starts at the second position of the 224th codon and includes 1,337 nucleotides of the 3' untranslated region.

The sequence of λ ADH5 (see Fig. 2) agrees with the predicted amino acid composition of the known human β_1 ADH subunit (Hempel *et al.* 1984). Comparison with the coding region of previously published β_1 cDNA clones shows only one difference at the third position of the 220th codon. The composite sequence of Heden *et al.* (1986) shows C, whereas λ ADH5, λ ADH8, and pUC ADH14 show T. Another possible difference is in the 236th codon of pUC

	(180)	(190)	(200)	630
	Tyr Gly Ser Ala Val Asn Val Ala Lys Val Thr Pro Gly Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly Leu Ser Ala Val Met			
λ5	TAT GGG TCT GCA GTT AAC GTT GCC AAG GTC ACC GCA GGC TCT ACC TGT GCT GTG TTT GGC CTG GGA GGC GTC GGC CTA TCT GCT GTT ATG			
λ8	---	---	---	---
p14	---	---	---	---
c	---	---	---	---
	(210)	(220)	(230)	720
	Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile Ala Val Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Leu Gly Ala Thr Glu			
λ5	GGC TGT AAA GCA GCT GGA GCA GCC AGA ATC ATT GCG GTG GAC ATC AAC AAG GAC AAA TTT GCA AAG GCC AAA GAG TTG GGT GCC ACT GAA			
λ8	---	---	---	---
λ11	---	---	---	---
p14	---	---	---	---
c	---	---	---	---
	(240)	(250)	(260)	810
	Cys Ile Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Lys Glu Met Thr Asp Gly Gly Val Asp Phe Ser Phe Glu Val Ile			
λ5	TGC ATC AAC CCT CAA GAC TAC AAG AAA CCC ATC CAG GAA GGC CTA AAG GAA GAC AAG GAT GAT GGA GGT GTG GAT TTT TCG TTT GAA GTC ATC			
λ8	---	---	---	---
λ11	---	---	---	---
p14	---	---	---	---
c	---	---	---	---
	(270)	(280)	(290)	900
	Gly Arg Leu Asp Thr Met Met Ala Ser Leu Leu Cys Cys His Glu Ala Cys Gly Thr Ser Val Ile Val Gly Val Pro Pro Ala Ser Glu			
λ5	GGT CGG CTT GAC ACC ATG ATG GCT TCC CTG TTA TGT TGT CAT GAG GCA TGT GGC ACA AGC GTC ATC GTA GGG GTA CCT CCT GCT TCC CAG			
λ8	---	---	---	---
λ11	---	---	---	---
p14	---	---	---	---
c	---	---	---	---
	(300)	(310)	(320)	990
	Asn Leu Ser Ile Asn Pro Met Leu Leu Leu Thr Gly Arg Thr Trp Lys Gly Ala Val Tyr Gly Gly Phe Lys Ser Lys Glu Gly Ile Pro			
λ5	AAC CTC TCA ATA AAC CCT ATG CTG CTA CTG ACT GGA CGC AGC TGG AAG GGG GGT GTT TAT GGT GGC TTT AAG AGT AAA GAA GGT ATC CGA			
λ8	---	---	---	---
λ11	---	---	---	---
p14	---	---	---	---
c	---	---	---	---
	(330)	(340)	(350)	1080
	Lys Leu Val Ala Asp Phe Met Ala Lys Lys Phe Ser Leu Asp Ala Leu Ile Thr His Val Leu Pro Phe Glu Lys Ile Asn Glu Gly Phe			
λ5	AAA CTT GTG GCT GAT TTT ATG GCT AAG AAG TTT TCA CTG GAT GCG TTA ATA ACC CAT GTT TTA CCT TTT GAA AAA ATA AAT GAA GGA TTT			
λ8	---	---	---	---
λ11	---	---	---	---
p14	---	---	---	---
c	---	---	---	---

Fig. 2-2

λ 11	AAAAACTGAA	CTATTCCCAG	AATCATGTTC	AAAAAATCTG	TAATTTTTCG	TGATGAAAGT	GCTTCATTGA	CTAAACAGTA	TTAGTTTGTG	GCTATAAATG	ATTATTTA	1944
c	-----	-----	-----	-----	-----*	-----	-----	-----	-----	-----	-----G-	
λ 11	ATGATGACTG	AAAATGTGTA	TAAAGTAATT	AAAAGTAATA	TGGTGGCCTT	AAGGTAGAG	ATGGGATGGC	AAATGCTGTG	AATGCAGAAT	GTAATAATTG	TAACATAAGAA	2054
c	-----	-----	-----*	-----	-----	-----	-----	-----	-----	-----	-----	
λ 11	ATGGCACAAA	CACGTTAAGC	AATATATTTT	CCTAGTAGAT	ATATATATAC	ACATACATAT	ATACACATAT	ACAAATGTAT	ATTTTGCAG	AATTGTTTTC	AATCTAGAAC	2164
c	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
λ 11	TTTTCTATTA	ACTACCATGT	CTTAAATCA	AGTCTATAAT	CCTAGCATT	GTTTAATATT	TTGAATATGT	AAAGACCTGT	GTTAATGCTT	TGTTAATGCT	TTTCCCACTC	2274
c	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
λ 11	TCATTTGTTA	ATGCTTTCCC	ACTCTCAGGG	GAAGGATTTG	CATTTTGAGC	TTTATCTCTA	AATGTGACAT	GCAAAGATTA	TTCTTGGTAA	AGGAGGTAGC	TGCTCTCCAA	2384
c	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
λ 11	AATGCTATTG	TTGCAATATC	TACATTCTAT	TTCATATTAT	GAAAGACCTT	AGACATAAAG	TAAATAGTT	TATCATT				2462
c	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	

Fig. 2-4

Fig. 2. Nucleotide sequence of cDNAs and deduced amino acid sequence for human ADH β_1 subunit (λ 5, λ 8, and λ 11). For comparison, the previously published sequences in Ikuta *et al.* (p14) and Heden *et al.* (c) are also shown. Numbers with and without parentheses are positions of amino acid residues from NH₂ terminus and nucleotide site, respectively. Initiation and termination codons are marked with +++ and ***, respectively. Deletion is denoted by *. Three arrows show the polyadenylation sites.

ADH14 (Ikuta *et al.* 1986), although the sequence GGG should probably be GGT as published previously for the same cDNA clone (Ikuta *et al.* 1985). Given this change in pUC ADH14, λ ADH5 and pUC ADH14 cDNA clones have sequences identical to the exons of the ADH2¹ gene published by Duester *et al.* (1986). These observations confirm an extremely low level of nucleotide diversity of the ADH β_1 subunit. In fact, the low level of nucleotide diversity has also been observed between two alleles which encode for ADHs β_1 and β_2 (Matsuo *et al.* 1987).

In the coding region, λ ADH8 shows two major differences from others. First, the second and third positions (CA) in the second codon (Thr) are deleted. Second, the second position of the 46th codon appears to be A rather than G, replacing Cys by Thr. If the deletion were real, then the 23rd residue would be the termination codon TGA, giving rise to an incomplete translation product. Furthermore, three residues 46 (Cys), 67 (His), and 174 (Cys) provide ligands to one catalytic zinc atom and four residues 97 (Cys), 100 (Cys), 103 (Cys), and 111 (Cys) provide ligands to another catalytic zinc atom (Ek-lund *et al.* 1976). All of these residues have been conserved perfectly at the amino acid level among class I ADHs from horse, human, rat, mouse, and maize (Yokoyama and Yokoyama 1987). Thus, the deletion of the two nucleotides and the single nucleotide substitution would drastically change the biological function of ADH. Therefore, these mutations of λ ADH8 are either due to a rare, incorrect transcription product or a cloning artifact. Despite these defects, this clone is still useful for evaluation of the level of DNA polymorphism and as an example of one of the three polyadenylation sites.

The clone λ ADH11 has an identical sequence to λ ADH5 and λ ADH8 over shared regions with one exception, a deletion of two nucleotides (GA) in the beginning of the 3' untranslated region. A comparison between λ ADH11 and the composite sequence of Heden *et al.* (1986) shows two nucleotide differences and four deletion-insertion polymorphisms, excluding the GA deletion in λ ADH11. Since all of these apparent deletion-insertions are contained in runs of several identical nucleotides, it is possible that they do not reflect true polymorphisms but may be sequencing or typographical errors. In summary, likely DNA polymorphisms are at the 1177th and 1313th nucleotides (Fig. 2), where pUC ADH14 and our three cDNA clones are identical but differ from that of Heden *et al.* (1986).

Heden *et al.* (1986) noted three different lengths of the 3' untranslated region: 213, 590, and 1330 nucleotides. The same three polyadenylation sites have been detected in our cDNA clones (Fig. 2). One minor difference occurs in the last position where both λ ADH6 and λ ADH11 contain the extra nucleotides TTT just before the poly(A) tail.

Restriction Fragment Length Polymorphism

A subclone of λ ADH11, called p1.1-1, contains approximately 700 nucleotides which starts just 3' to the second polyadenylation site through the third polyadenylation tail (see Fig. 1). This single-stranded subclone was labelled with ^{32}P according to Jeffreys *et al.* (1985) and hybridized to Rsa I digested DNA from a total of 27 individuals: 16 U.S. white, 5 non-U.S. blacks, 4 Japanese, 1 Indian and 1 U.S. black. Fig. 3 shows 8 of the 27 samples. All

Fig. 3. Southern hybridization of human genomic DNA with probe p1.1-1 (see Fig. 1). Approximately 10 μg of human genomic DNA from eight different individuals (a-h) were digested with the restriction enzyme Rsa I, separated on a 1% agarose gel, blotted and hybridized to the ^{32}P -labelled probe, p1.1-1. Lanes a-e contain DNA from five non-U.S. blacks (Nigeria, Ivory Coast, Zaire, Papua New Guinea, Zambia, respectively), f from a U.S. black, g and h from two Japanese individuals. DNAs of the non-U.S. blacks were kindly provided by Dr. Satoshi Horai. The sizes (in kilobases) of the hybridizing bands are indicated on the right.

lanes except lane f are representative of the 26 out of the 27 individuals sampled by having two bands of sizes 3.6 and 1.3 kb. From partially characterized genomic clones, it is known that the 3.6 kb and 1.3 kb fragments correspond to ADH2 and ADH1 loci, respectively (Matsuo and Yokoyama, unpublished observations). Lane f shows a different banding pattern, which consists of three fragments of sizes 3.7, 2.2 and 1.3 kb. It appears that this individual has the invariant 1.3 kb fragment at the ADH1 locus but is most likely heterozygous at the ADH2 locus, both alleles being different from the other 26 individuals.

Molecular Evolution of Class I ADH

To clarify the evolutionary relationship and rates of nucleotide substitution of α , β , γ subunits, we first constructed a phylogenetic tree of the three subunits and then evaluated the different branch lengths. The nucleotide sequences of cDNAs α and γ are from Ikuta *et al.* (1986) and that of cDNA β is a composite from pUC ADH14 and λ ADH5, λ ADH8, and λ ADH11. To evaluate all of the branch lengths, the nucleotide sequence of the cDNA from mouse (Edenberg *et al.* 1985) is used as a reference.

Proportions of observed nucleotide differences among the four sets of nucleotide sequences (p, above the diagonal) and the corresponding numbers of nucleotide substitutions per nucleotide site (d, below the diagonal) are shown in Table 1. The latter quantity is obtained from the relationship $d = -(3/4) \ln [1 - (4/3)p]$, which assumes equal substitution rates among the four types of nucleotides (Jukes and Canter 1969). By using the transformed distance method (Nei 1987), the d values are transformed as follows: $d_{\alpha\beta} = (0.0507 - 0.2165 - 0.2082)/2 + 0.2094 = 0.0224$; $d_{\alpha\gamma} = (0.0651 - 0.2165 - 0.2035)/2 + 0.2094 = 0.0320$; $d_{\beta\gamma} = (0.0412 - 0.2082 - 0.2035)/2 + 0.2094 = 0.0242$. Here, the value 0.2094 is an arithmetic mean of three values 0.2165, 0.2082 and 0.2035 (see Table 1), which is introduced to make $d_{\alpha\beta}$, $d_{\alpha\gamma}$, and $d_{\beta\gamma}$ positive. The computation shows that the subunits α and β diverged most recently and that

Table 1. Proportions of nucleotide differences (above the diagonal) and the corresponding numbers of nucleotide substitutions per site (below the diagonal) for human ADHs α , β , and γ , and mouse ADH

	ADH α	ADH β	ADH γ	mouse ADH
ADH α		55/1122 (0.0490)	70/1122 (0.0624)	211/1122 (0.1881)
ADH β	0.0507 (± 0.0069)		45/1122 (0.0401)	204/1122 (0.1818)
ADH γ	0.0651 (± 0.0079)	0.0412 (± 0.0062)		200/1122 (0.1783)
mouse ADH	0.2165 (± 0.0156)	0.2082 (± 0.0152)	0.2035 (± 0.0150)

Fig. 4. A phylogenetic tree for the α , β , and γ subunits, where A and B denote branch points. Numbers besides different branches are lengths estimated by the method of Fitch and Margoliash (1967).

their common ancestor diverged from that of the subunit γ earlier (Fig. 4).

Proportions of observed nucleotide differences between α and β , between α and γ , and between β and γ are 0.0490, 0.0624, and 0.0401, respectively (see Table 1). Thus, if the evolutionary rates of the three subunits were equal, then the data could be interpreted such that the subunits β and γ are most closely related and diverged most recently. In fact, Ikuta *et al.* (1986) recently proposed such an evolutionary scheme, which, however, is not supported from the present analyses. Further support for our conclusion can be reached by comparing amino acid sequences of ADHs from horse, human, mouse, rat, and maize (Yokoyama and Yokoyama 1987). The fast evolution of the subunit α can be clearly seen in the comparison of nucleotide sequences of α , β , and γ to the reference, i.e., mouse ADH. When compared to the mouse sequence, the frequencies of nucleotide differences of the subunits α , β , and γ are given by 0.1881, 0.1818, and 0.1783, respectively (Table 1). When evolutionary rates vary from one lineage to another, the transformed distance method gives an improved topology (Nei 1987).

To estimate the lengths of different branches in Fig. 4 without making a strong assumption of rate homogeneity, the method of Fitch & Margoliash (1967) is used. Then, the lengths of the branches A-B, B- α , B- β , and A- γ are

estimated to be 0.0057 ± 0.0004 , 0.0334 ± 0.0055 , 0.0173 ± 0.0040 , and 0.0222 ± 0.0045 , respectively. The total lengths of the subunits α , β , and γ from the branch point A are given by 0.0391, 0.0230, and 0.0222, respectively (Fig. 4). Thus, subunit γ is evolving with the slowest rate, followed by the subunits β and α , in that order (see also Yokoyama and Yokoyama 1987). The differences in the total lengths between A- α and A- γ and between B- α and B- β are significantly different, whereas that between A- β and A- γ is not.

4. DISCUSSION

In the studies of the human ADHs, the subunit β has received much attention, partly because of the high frequency of acute alcohol intoxication (Wolff 1972) and the high frequency of the β_2 subunit (Stamatoyannopoulos *et al.* 1975) in Orientals. Furthermore, genetic polymorphism at the ADH2 locus is suspected to have importance in determining the degree of susceptibility to fetal alcohol syndromes (Streissguth *et al.* 1980), alcoholism (Li 1977), and chlorpropamide-induced alcohol sensitivity (Leslie and Pyke 1978) (for a review, see Yoshida 1982). With that in mind, we were interested in 1) comparing nucleotide sequences of published ADH2 for nucleotide diversity, 2) finding a probe which can detect restriction length polymorphisms associated with ADH2 and 3) determining a phylogenetic tree for the three class I human ADHs.

Nucleotide differences of the presented ADH2¹ clones with previously published ADH2¹ clones (Ikuta *et al.* 1986; Heden *et al.* 1986) shows an extremely low level of nucleotide diversity: one silent nucleotide substitution in the coding region and two nucleotide substitutions in the 3' untranslated region.

Restriction fragment length polymorphism (RFLP) studies have been particularly useful in conducting population studies and detecting differences associated with genetic diseases. The ADH2 3' region (p1.1-1) described in the present paper, seems to be a useful probe for detecting the locus specific Rsa I polymorphism (as shown in Fig. 3). Since the one individual which showed a different hybridizing pattern from the others was the only U.S. black sampled, it will be interesting to see if this is a common RFLP to be found in U.S. blacks.

The evolutionary relationships between the α , β and γ subunits derived from the cDNA sequences and amino acid comparison (Yokoyama and Yokoyama 1987) seem to show that α has evolved fastest and γ evolved slowest. It has been observed that the isozyme $\gamma\gamma$ oxidizes ethanol much better than $\beta\beta$ (Wagner *et al.* 1983; Buhler and Von Wartburg 1982; Bosron *et al.* 1983). The slow rate of nucleotide substitution in the γ subunit is consistent with one of the characteristics of molecular evolution: "Functionally less important molecules or parts of molecules evolve (in terms of mutant substitutions)

faster than more important ones (Kimura 1983)". These observations seem to suggest that the subunit γ is providing the original function in the ethanol metabolism. It also suggests that ADH γ may be more important in ethanol metabolism than ADH β , and that genetic polymorphism at the ADH3 locus may be more significant in determining the degree of susceptibility to various alcohol-related syndromes.

We are grateful to S. L. C. Woo and T. Chandra for supplying the human liver cDNA library. This work was supported by U.S. Public Health Service Grants GM-28672 and AA-03539 (S.Y.) and Postdoctoral Training Grant MH-14677 (R.Y.). P.R. acknowledges a Career Development Award from the Juvenile Diabetes Foundation. Part of this work was done while S.Y. and R.Y. were affiliated with Washington University.

REFERENCES

- BIGGIN, M. D., GIBSON, T. J. and HONG, G. F. (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* **80**, 3963-3965.
- BOSRON, W. F., MAGNAS, L. J. and LI, T. -K. (1983) Kinetic and electrophoretic properties of native and recombined isozymes of human liver alcohol dehydrogenase. *Biochemistry* **22**, 1852-1857.
- BUHLER, R., HEMPEL, J., KAISER, R., DE ZALENSKI, C., VON WARTBERG, J. -P. and JORNVALL, H. (1984) Human liver alcohol dehydrogenase. 2. The primary structure of the γ_1 protein chain. *Eur. J. Biochem* **145**, 447-453.
- BUHLER, R. and VON WARTBURG, J. -P. (1982) Purification and substrate specificities of three human liver alcohol dehydrogenase isoenzymes. *FEBS Lett* **144**, 135-139.
- DUESTER, G., HATFIELD, G. W., BUHLER, R., HEMPEL, J., JORNVALL, H. and SMITH, M. (1984) Molecular cloning and characterization of a cDNA for the β subunit of human alcohol dehydrogenase. *Proc. Natl. Acad. Sci. USA* **81**, 4055-4059.
- DUESTER, G., SMITH, M., BILANCHONE, V. and HATFIELD, G. W. (1986) Molecular analysis of the human class I alcohol dehydrogenase gene family and nucleotide sequence of the gene encoding the β subunit. *J. Biol. Chem.* **261**, 2027-2033.
- EDENBERG, H. J., ZHANG, K., FONG, K., BOSRON, W. F. and LI, T. -K. (1985) Cloning and sequencing of cDNA encoding the complete mouse liver alcohol dehydrogenase. *Proc. Natl. Acad. Sci. USA* **82**, 2262-2266.
- EKLUND, H., NORDSTROM, B., ZEPPEAUER, E., SODERLUND, G., OHLSSON, I., BOINE, T., SODERBERG, B. O., TAPIA, O., BRANDEN, C. -I. and AKESON, A. (1976) Three dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution. *J. Mol. Biol.* **102**, 27-59.
- FITCH, W. F. and MARGOLIASH, E. (1967) Construction of phylogenetic trees. *Science* **155**, 279-284.
- HARADA, S., MISAWA, S., AGARWAL, D. P. and GOEDDE, H. W. (1980) Liver alcohol dehydrogenase and aldehyde dehydrogenase in Japanese: Isozyme variation and its possible role in alcohol intoxication. *Am. J. Hum. Genet.* **32**, 8-15.
- HEDEN, L. -O., HOOG, J. -O., LARSSON, K., LAKE, M., LAGERHOLM, E., HOLMGREN, A., VALLEE, B. L., JORNVALL, H. and VON BAHR-LINDSTROM, H. (1986) cDNA clones coding for the β -subunit of human liver alcohol dehydrogenase have differently sized 3'-non-coding regions. *FEBS Lett* **194**, 327-332.
- HELMS, C., GRAHAM, M. Y., DUTCHIK, J. E. and OLSON, M. Y. (1985) A new method for purifying lambda DNA from phage lysates. *DNA* **4**, 39-49.
- HEMPEL, J., BUHLER, R., KAISER, R., HOLMQUIST, B., DE ZALENSKI, C., VON WARTBURG, J. -P., VALLEE, B. and JORNVALL, H. (1984) Human liver alcohol dehydrogenase. 1. The primary structure of the $\beta_1\beta_1$ isozyme. *Eur. J. Biochem.* **145**, 437-445.
- IKUTA, T., SZETO, S. and YOSHIDA, A. (1986) Three human alcohol dehydrogenase subunits: cDNA

- structure and molecular and evolutionary divergence. *Proc. Natl. Acad. Sci. USA* **83**, 634-638.
- JEFFREYS, A. J., WILSON, V. and THEIN, S. L. (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**, 67-73.
- JUKES, T. H. and CANTOR, C. R. (1969) Evolution of protein molecules. In *Mammalian Protein Metabolism III*. (ed. Munro, H. N.) pp. 21-132. Academic Press, New York.
- KAN, Y. W. and DOZY, A. M. (1978) Polymorphism of DNA sequence adjacent to human β -globin structural gene: Relationship to sickle mutation. *Proc. Natl. Acad. Sci. USA* **75**, 5631-5635.
- KIMURA, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- LARSON, R. and MESSING, J. (1983) Apple II computer for DNA and protein sequence data. *DNA* **2**, 31-35.
- LESLIE, R. D. G. and PYKE, D. A. (1978) Chlorpropamide-alcohol flushing: a dominantly inherited trait associated with diabetes. *Brit. Med. J.* **2**, 1519-1521.
- LI, T. -K. (1977) Enzymology of human alcohol metabolism. In *Alcoholism*. pp. 19-56. John Wiley and Sons, New York.
- MANIATIS, T., FRITSCH, E. F. and SAMBROOK, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor New York.
- MATSUO, Y., YOKOYAMA, R. and YOKOYAMA, S. (1987) Human alcohol dehydrogenases β_1 and β_2 are characterized by a single nucleotide substitution. *Science* (submitted)
- NEI, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- NORRANDER, J., KEMPE, T. and MESSING, J. (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* **26**, 101-106.
- SANGER, F., NICKLEN, S. and COULSEN, A. R. (1977) DNA sequencing with chain-terminating inhibitor. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- SMITH, M. (1986) Genetics of human alcohol and aldehyde dehydrogenases. *Adv. Hum. Genet.* **14**, 249-290.
- SMITH, M., HOPKINSON, D. A. and HARRIS, H. (1972) Alcohol dehydrogenase isozymes in stomach and liver: Evidence for activity of the ADH3 locus. *Ann. Hum. Genet.* **35**, 243-253.
- SMITH, M., HOPKINSON, D. A. and HARRIS, H. (1973a) Studies on the properties of the human alcohol dehydrogenase isozymes determined by the different loci ADH1, ADH2, ADH3. *Ann. Hum. Genet.* **37**, 49-67.
- SMITH, M., HOPKINSON, D. A. and HARRIS, H. (1973b) Studies on the subunit structure and molecular size of the human alcohol dehydrogenase isozymes determined by the different loci ADH1, ADH2, ADH3. *Ann. Hum. Genet.* **36**, 401-414.
- SOUTHERN, E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.
- STAMATOYANNOPOULOS, G., CHEN, S. -H. and FUKUI, M. (1975) Liver alcohol dehydrogenase in Japanese: High frequency of atypical form and its role in alcohol sensitivity. *Am. J. Hum. Genet.* **27**, 789-796.
- STREISSGUTH, A. P., LANDESMAN-DWYER, S., MARTIN, J. C. and SMITH, D. W. (1980) Teratogenic effects of alcohol in humans and laboratory animals. *Science* **209**, 353-361.
- TENG, Y. S., JEHAN, S. and LIE-IJO, L. E. (1979) Human alcohol dehydrogenase ADH₂ and ADH₃ polymorphisms in ethnic Chinese and Indians of West Malaysia. *Hum. Genet.* **53**, 87-90.
- VON BAHR-LINDSTROM, H., HOOG, J. -O., HEDEN, L. -O., KAISER, R., FLEETWOOD, L., LARSSON, K., LAKE, M., HOLMQUIST, B., HOLMGREN, A., HEMPEL, J., VALLEE, B. L. and JORNVALL, H. (1986) cDNA and protein structure for the α subunit of human liver alcohol dehydrogenase. *Biochemistry* **25**, 2465-2470.
- WAGNER, F. W., BURGER, A. R. and VALLEE, B. L. (1983) Kinetic properties of human liver alcohol dehydrogenase: Oxidation of alcohols by class I isozymes. *Biochemistry* **22**, 1857-1863.
- WATANABE, S., KONDO, S. and MATSUNAGA, E. (1975) *Anthropological and Genetic Studies on the Japanese*. University of Tokyo Press, Tokyo.
- WILBUR, W. J. and LIPMAN, D. J. (1983) Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* **80**, 726-730.

- WOLFF, P. H. (1972) Ethnic differences in alcohol sensitivity. *Science* **175**, 449-450.
- Woo, S. L. C. (1979) A sensitive and rapid method for recombinant phage screening. *Meth. Enzymol.* **68**, 389-395.
- YOKOYAMA, S. and YOKOYAMA, R. (1987) Molecular evolution of mammalian class I alcohol dehydrogenases. *Mol. Biol. Evol.*, in press.
- YOSHIDA, A. (1982) Molecular basis of difference in alcohol metabolism between Orientals and Caucasians. *Jpn. J. Hum. Genet.* **27**, 55-70.
- YOUNG, R. A. and DAVIS, R. W. (1983) Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA* **80**, 1194-1198.