Molecular Evolution of the Zinc-containing Long-Chain Alcohol Dehydrogenase Genes[1]

Shozo Yokoyama, Ruth Yokoyama,* Claire S. Kinlaw,‡ and David E. Harry†*

*Department of Ecology, Ethology, and Evolution and †Department of Forestry, University of Illinois at Urbana-Champaign; and ‡Institute of Forest Genetics, USDA Forest Service at Berkeley

Phylogenetic relationships and rates of nucleotide substitution were studied for alcohol dehydrogenase (ADH) genes by using DNA sequences from mammals and plants. Mammalian ADH sequences include the three class I genes and a class II gene from humans and one gene each from baboon, rat, and mouse. Plant sequences include two ADH genes each from maize and rice, three genes from barley, and one gene each from wheat and two dicots, *Arabidopsis* and pea. Phylogenetic trees show that relationships among ADH genes are generally consistent with taxonomic relationships: mammalian and plant ADH genes are classified into two distinct groups; primate class I genes are clustered; and two dicot sequences are clustered separately from monocot sequences. Accelerated evolution has been detected among the duplicated ADH genes in plants, in which synonymous substitutions occurred more often within the coenzyme-binding domain than within the catalytic domains.

Introduction

Alcohol dehydrogenases (ADHs) constitute two evolutionary groups, one characterized by short protein chains (~250 residues) which do not require zinc as a cofactor and one characterized by long protein chains (~370 residues) which require zinc as a cofactor (Jörnvall et al. 1981, 1984). The former group is represented by ADHs from *Drosophila,* and the latter is represented by ADHs from organisms as diverse as mammals, plants, and yeasts (Jörnvall 1985). These two ADH groups are so distantly related that the alignment of their amino acid sequences is virtually impossible (also see Benyajati et al. 1981). However, the amino acid sequences of the zinc-containing long-chain (LC) ADHs are highly conserved. For example, maize and mammalian ADHs have an amino acid sequence similarity of ~50%, and mammalian and budding yeast ADHs share ~20% similarity (Eklund et al. 1976*a;* Jörnvall et al. 1987).

All of the LC ADHs are dimeric enzymes. In humans three classes (I, II, and III) of ADH are distinguished by their ability to form intergenic heterodimers. In class I, three subunits—α, β, and γ—are encoded by three distinct loci: ADH1, ADH2, and ADH3, respectively. These subunits are randomly combined to produce dimeric enzymes. Class II subunits (π) and class III subunits (χ) are encoded by ADH4 and

1. Key words: alcohol dehydrogenase, molecular evolution, phylogeny.

Address for correspondence and reprints: Dr. Shozo Yokoyama, Department of Ecology, Ethology, and Evolution, University of Illinois, 505 South Goodwin Avenue, Urbana, Illinois 61801.

Mol. Biol. Evol. 7(2):143–154. 1990.
© 1990 by The University of Chicago. All rights reserved.
0737-4038/90/0702-0003$02.00

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1012357 by guest on 20 August 2022

ADH5, respectively. As in humans, murine ADHs are also subdivided into the three classes (e.g., see Edenberg et al. 1985; Julia et al. 1988). Among these, class I ADHs comprise the major isozymes involved in ethanol metabolism (for a review, see Smith 1986). ADHs from plants have not been similarly classified, but they are also dimeric and in many instances readily form intergenic heterodimers (e.g., see Harberd and Edwards 1983; Freeling and Bennett 1985).

To date, most evolutionary comparisons of LC ADHs have been based on amino acid sequences (e.g., see Eklund et al. 1976*a;* Ikuta et al. 1986; Jörnvall et al. 1987; Yokoyama and Yokoyama 1987). However, because the genetic code is degenerate, evolutionary comparisons based on DNA sequences are more informative than those based on amino acid sequences. Since nucleotide sequences of ADH cDNA and genomic clones from mammals and plants have recently become available, such analyses are now feasible. In the present paper, we construct phylogenetic trees of LC ADH genes from mammals and plants and study the patterns of nucleotide substitution in different functional domains of the molecule during different evolutionary stages.

Material and Methods

Class I ADH genes in humans (*Homo sapiens*) were represented by alleles from ADH1 (denoted HS1; Ikuta et al. 1986; Matsuo and Yokoyama 1989), ADH2 (HS2; Hempel et al. 1984; Ikuta et al. 1986; Yokoyama et al. 1987), and ADH3 (HS3; Bühler et al. 1984; Ikuta et al. 1986) loci. It should be noted that the ADH1 locus is nearly monomorphic (Matsuo and Yokoyama 1989) and that nucleotide variation at the ADH2 (Ehrig et al. 1988; Matsuo et al. 1989) and ADH3 (Höög et al. 1986) loci is low. Thus, we included sequence data for only one allele from each locus. Specifically, HS2 is represented by an allele ($ADH2^1$) that encodes protein subunit $ADH\beta_1$, and HS3 is represented by an allele ($ADH3^1$) that encodes subunit $ADH\gamma_1$. Furthermore, class I ADH genes from baboon [*Papio papio* (PP); Trezise et al. 1989], rat [*Rattus norvegicus* (RN); Crabb and Edenberg 1986], and mouse [*Mus musculus* (MM); Edenberg et al. 1985] and the class II ADH gene from human (HS4; Höög et al. 1987) were included.

For plants, ADH genes from four monocots and two dicots were included. ADH1 from maize [*Zea mays* (ZM1)] was represented by the ADH1-F allele (Dennis et al. 1984), whose sequence differs only slightly from that of a second allele, ADH1-S (Sachs et al. 1986). Maize was also represented by a second gene, ADH2 (ZM2; Dennis et al. 1985). ADH genes from barley (*Hordeum vulgare*) included ADH1 (HV1; Good et al. 1988; Trick et al. 1988), ADH2 (HV2; Trick et al. 1988), and ADH3 (HV3; Trick et al. 1988). An ADH gene (TA) from wheat *Triticum aestivum* (Mitchell et al. 1989) and ADH1 (OS1) and ADH2 (OS2) genes from rice [*Oryza sativa* (Xie and Wu 1989)] were also included.

Dicot sequences used were those of *Arabidopsis thaliana* (AT; Chang and Meyerowitz 1986) and pea [*Pisum sativum* (PS); Llewellyn et al. 1987] ADH genes.

The number of amino acid residues and their sequences are highly conserved among these LC ADHs. When the initiation and stop codons are excluded from consideration, HS1, HS2, HS3, PP, and MM each have 374 codons, while RN and HS4 have 375 and 379 codons, respectively. Among plants, HV1, HV3, ZM1, ZM2, TA, and AT each consist of 379 codons, while PS, OS1, and OS2 have 380, 376, and 375 codons, respectively. HV2 is somewhat unusual and has 373 codons, lacking six codons between codons 293 and 298 when it is compared with others (Trick et al. 1988).

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1013... by guest on 20 August 2022

Because of the small number of deletions/insertions, ADH sequences can be easily aligned visually (e.g., see Yokoyama and Yokoyama 1987, fig. 1). To estimate the number of nucleotide substitutions, a total of 368 codons were used.

From a study of X-ray crystallography, a three-dimensional structure of horse ADH has been proposed, and functional domains have been identified (Eklund et al. 1976*b*). When HS1 is used as a reference, one catalytic domain (CD1) spans codons 1–175; the coenzyme binding domain (CBD) spans codons 176–318; and a second catalytic domain (CD2) spans codons 319–374. Because the rate and pattern of amino acid replacements differ among these domains in mammals (Yokoyama and Yokoyama 1987), they were distinguished in the analyses. Furthermore, the distinction of the three domains allows us to study whether selective constraints differ among different functional domains.

Pairwise comparisons were made for all 17 ADH genes by calculating the number of nucleotide substitutions per site (d). These values were estimated from $d = -(3/4)\ln[1-(4/3)p]$, where p is the proportion of observed nucleotide differences between two sequences (Jukes and Cantor 1969). We also computed the d values by using other formulas (Kimura 1981; Takahata and Kimura 1981; Gojobori et al. 1982), but the values obtained were very similar. This is reasonable because the d values for these genes are considerably smaller than one (table 1; and see Gojobori et al. 1982; Tajima and Nei 1984).

In addition to the total d, it is useful to distinguish the number of nucleotide substitutions that do not produce an amino acid replacement [synonymous substitutions (d_s)] from the number of those that do produce amino acid replacement [nonsynonymous substitutions (d_n)]. d_s and d_n may vary among different lineages, affecting evolutionary interpretations. To study such possibilities, d_s and d_n values were evaluated by using the method of Miyata and Yasunaga (1980).

Topology and branch lengths of phylogenetic trees were evaluated by using both the neighbor-joining (NJ) method (Saitou and Nei 1987) and the unweighted pair-group method with arithmetic mean (UPGMA) (Sokal and Michener 1958; for details of the procedure, see Nei 1987, pp. 293–298).

Results

d and d_n Values

Estimates of d and d_n for the entire coding region are given in table 1. Comparisons within mammalian class I ADH genes show that $d < 0.3$, whereas differences within plant ADH genes show $d < 0.4$ (table 1). Even when HS4 is included, d values among mammalian ADH genes are <0.6. However, differences between plant and animal genes are larger, showing $d > 0.6$ (table 1).

A Phylogenetic Tree

Phylogenetic trees for the ADH genes were constructed by using the NJ method based on d (fig. 1A) and d_n (fig. 1B) values for the entire coding sequence. Except for minor differences in the relationship among the monocot ADH genes, the two trees have similar topologies. Phylogenetic trees constructed by the UPGMA method also gave similar tree topologies. As expected, plant ADH genes form one cluster and mammalian ADH genes form another.

Among mammalian ADH genes, HS1, HS2, HS3, and PP are most closely related. The common ancestor of these class I genes diverged from that of MM and RN, which

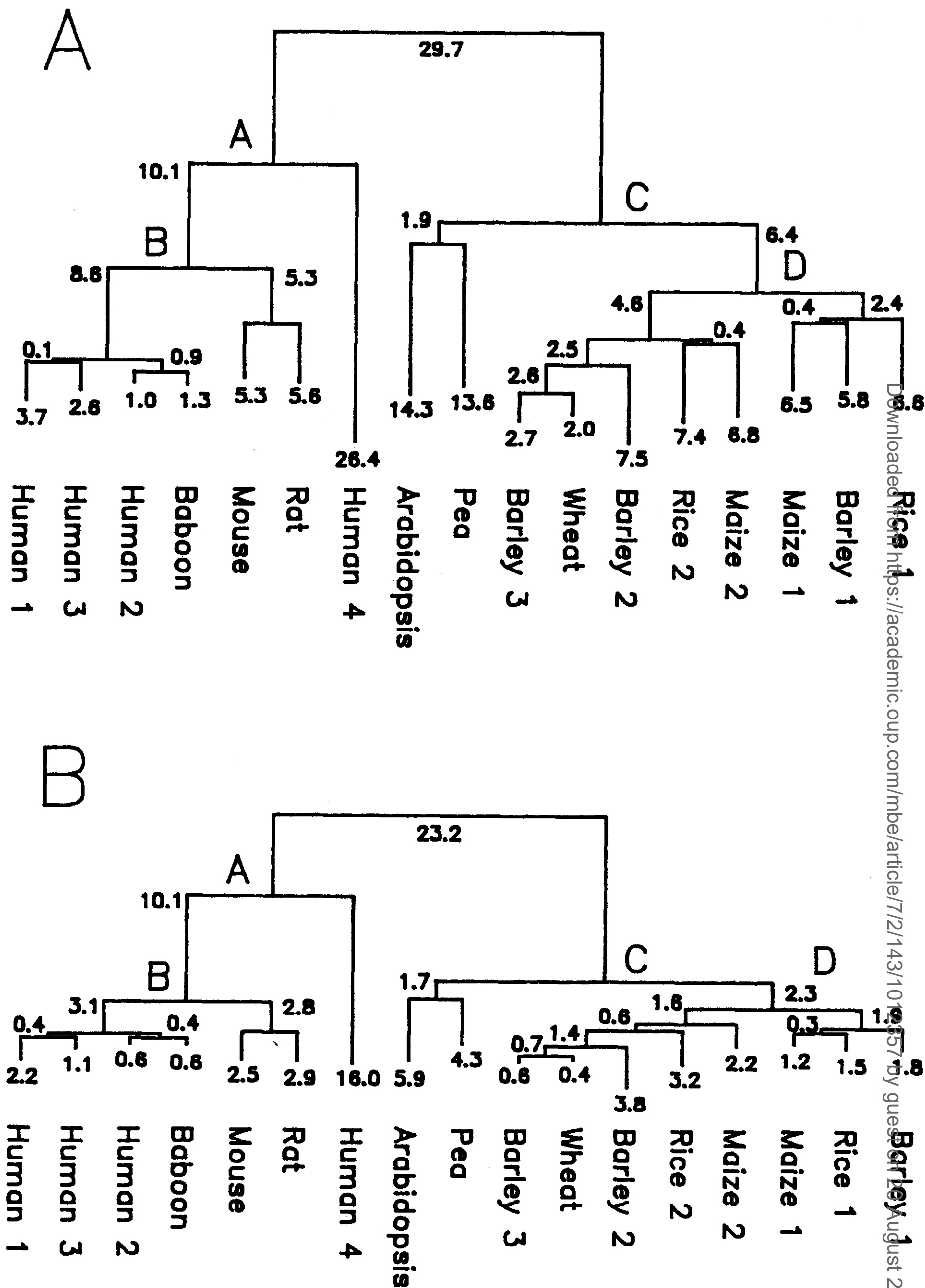
Downloaded from https://academic.oup.com/mbe/article/7/2/143/1012357 by guest on 20 August 2022

Table 1
Pairwise Comparisons of ADH Genes

	HS1	HS2	HS3	PP	MM	RN	HS4	AT	PS	ZM1	ZM2	HV1	HV2	HV3	TA	OS1	OS2
HS1		5.1	6.3	6.4	21.5	24.2	44.7	71.2	69.4	68.4	68.0	67.1	74.8	71.4	70.7	70.7	69.7
HS2	3.1		4.4	2.3	20.6	22.5	45.5	70.8	68.1	65.4	66.0	65.0	71.2	70.0	68.2	68.4	66.6
HS3	3.3	2.6		5.4	20.0	22.0	45.0	71.2	68.3	66.9	67.6	66.3	73.8	70.2	69.5	68.2	68.8
PP	3.9	1.3	3.2		21.7	23.7	45.7	70.1	66.5	64.8	66.9	64.1	71.9	70.5	68.2	68.4	67.0
MM	9.8	9.6	8.7	10.0		10.9	48.9	67.4	69.9	67.1	66.0	67.8	70.5	66.3	66.0	68.9	67.0
RN	11.3	10.6	10.2	11.4	5.4		51.1	67.8	69.2	69.8	66.3	66.9	69.8	66.0	65.8	67.8	67.4
HS4	30.8	30.0	30.2	30.1	31.7	32.4		70.2	69.5	71.3	74.6	72.4	77.7	76.0	74.8	70.8	77.6
AT	45.9	43.6	44.4	44.0	45.0	44.6	45.4		27.9	32.9	34.7	31.4	36.4	35.4	35.1	31.6	34.1
PS	47.6	45.4	46.5	44.7	46.8	46.4	45.8	10.2		31.0	36.1	28.8	38.8	34.8	33.0	29.0	34.4
ZM1 . . .	45.2	43.7	43.9	43.4	45.2	45.2	45.1	13.4	11.0		20.3	13.5	25.0	21.7	20.2	12.3	20.7
ZM2 . . .	44.5	43.4	43.7	43.1	44.3	44.8	45.4	14.5	12.4	6.9		21.0	16.5	16.1	15.8	20.9	14.2
HV1 . . .	45.7	44.0	44.2	43.4	45.1	44.9	46.7	13.3	11.5	2.7	7.0		24.1	20.6	20.4	12.7	23.0
HV2 . . .	47.6	45.7	46.7	45.8	47.4	47.1	47.1	17.1	15.6	10.8	8.4	10.6		12.3	12.6	23.1	16.9
HV3 . . .	45.4	44.4	44.6	44.1	45.4	45.3	46.2	15.5	14.0	8.2	5.7	8.4	5.0		4.7	20.8	15.9
TA	44.8	43.6	43.9	43.3	44.8	44.5	45.4	15.3	13.4	7.7	5.5	8.0	5.1	1.1		19.5	15.3
OS1	45.8	44.2	44.9	44.4	45.6	46.1	45.9	12.7	10.3	3.0	7.6	3.8	11.0	8.6	8.1		21.1
OS2	47.6	45.7	46.3	45.5	46.6	47.3	47.4	15.2	13.7	9.3	5.7	9.8	8.4	5.8	5.8	9.1	

NOTE.—Values above and below the diagonal are $d \times 100$ and $d_n \times 100$, respectively. The number of nucleotide sites used for estimating d and d_n were 1,104 and 849, respectively.

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1012357



Downloaded from https://academic.oup.com/mbe/article/7/2/143/1016557 by guest on 21 August 2022

FIG. 1.—Phylogenetic trees constructed by considering d in the entire region of the DNA (A) and that of nonsynonymous substitutions (B). The numbers next to the different branches are branch lengths in 100 times the nucleotide substitutions per site, as estimated by means of the NJ method of Saitou and Nei (1987). In the text, Human 1, Human 2, Human 3, Human 4, Baboon, Mouse, Rat, Arabidopsis, Pea, Barley 1, Barley 2, Barley 3, Rice 1, Rice 2, Wheat, Maize 1, and Maize 2 are denoted as HS1, HS2, HS3, HS4, PP, MM, RN, AT, PS, HV1, HV2, HV3, OS1, OS2, TA, ZM1, and ZM2, respectively.

in turn earlier diverged from the ancestor of HS4. Among plant ADH genes, the common ancestor of HV2, HV3, TA, OS2, and ZM2 diverged from that of OS1, HV1 and ZM1. This ancestral monocot ADH gene had earlier diverged from the dicot ancestral gene. These phylogenetic relationships among the monocot genes are consistent with the observation that ZM1 and HV1 encode subunits that reassociate in vitro to form functional heterodimers (Trick et al. 1988) and that are the predominant isozymes induced by anaerobic stress (Harberd and Edwards 1983; Freeling and Bennett 1985; Xie and Wu 1989).

Branch Lengths

d Values

Figure 1A shows that HS4 is evolving faster than other mammalian genes. Differences in lengths between A–HS4 and other comparable branches of mammalian genes are all statistically significant, with the exception of the comparison between A–HS4 and A–HS1. Some representative branch lengths are shown in table 2, where lengths between CD1 and CBD and between CD2 and CBD have been compared. By comparing branches A–HS4 and A–HS2, we can see that the accelerated evolution of HS4 is due to faster changes in both CD1 and CBD regions in that gene (table 2). Table 2 also shows that CD2 regions of murine genes are evolving more slowly than are CD1 and CBD regions of murine genes and CD2 regions of primate genes.

Among the plant ADH genes, HV2, HV3, and OS2 are evolving faster than HV1 and OS1 (see fig. 1A); that is, branches D–HV2, D–HV3, and D–OS2 are significantly longer than D–HV1 and D–OS1. Branch C–HV2 is also significantly longer than C–AT and C–PS, showing that HV2 is evolving faster than AT and PS. Comparison of

Table 2
Branch Lengths Measured by *d* Values × 10^2

Branch	Domain			All Domains
	CD1 (174 codons)	CBD (139 codons)	CD2 (55 codons)	(368 codons)
Mammals:				
A–HS4	28.8 ± 2.7	25.1 ± 2.8	21.9 ± 4.1	26.4 ± 1.8
A–HS2	21.4 ± 2.3	18.9 ± 2.3	25.4 ± 4.5	20.6 ± 1.0
B–HS1	13.8 ± 1.7	10.0 ± 1.6	15.5 ± 3.3	12.4 ± 1.1
B–HS2	10.7 ± 1.5	10.9 ± 1.7	10.2 ± 2.6	10.5 ± 1.0
B–MM	12.3 ± 1.6	9.6 ± 1.6	6.8 ± 2.1	10.6 ± 1.0
B–RN	12.3 ± 1.6	10.9 ± 1.7[a]	5.5 ± 1.9[a]	10.9 ± 1.0
Plants:				
C–AT	15.8 ± 1.9	16.4 ± 2.2	13.5 ± 3.1	16.2 ± 1.3
C–PS	14.0 ± 1.8	15.2 ± 2.1	19.4 ± 3.8	15.5 ± 1.3
D–HV2 ...	11.7 ± 1.6[b]	19.4 ± 2.4[b]	15.6 ± 3.3	14.6 ± 1.2
D–OS2	11.2 ± 1.6	16.2 ± 2.1[c]	8.6 ± 2.4[c]	12.4 ± 1.1
D–ZM1 ...	7.6 ± 1.2	11.3 ± 1.7	9.5 ± 2.5	9.3 ± 1.0

NOTE.—See fig. 1 for branch points A, B, C, and D. Standard errors were computed from $[9p(1-p)/\{(3-4p)^2n\}]^{1/2}$, where p is the proportion of sites that differ and n is the number of nucleotide sites involved (see Nei 1987).

[a] Difference in branch length is significant at the 5% level.

[b] Difference in branch length is significant at the 1% level.

[c] Difference in branch lengths is significant at the 5% level.

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1123357 by guest on 20 August 2022

the three functional domains among plant genes shows that CBD is frequently evolving more rapidly than CD1 and CD2 (table 2).

d_n Values

When only d_n was considered, branch lengths from A to the various mammalian ADH genes were more uniform than when total d was considered (see fig. 1B). In particular, HS4 is evolving at a rate similar to that of the other mammalian ADH genes. The difference in the branch lengths between A–HS4 and the comparable branches of mammalian genes in figure 1A is caused by the difference in d_s.

Yokoyama and Yokoyama (1987) observed that rates of amino acid replacement in mouse and rat ADHs are significantly higher than in human and horse ADHs. The same trend can be seen in figure 1B: murine ADH genes are evolving slightly faster than HS2, HS3, and PP (table 3 and fig. 1B), but the differences in the branch lengths between the murine and primate ADH genes are not statistically significant. The rate of amino acid replacement and the rate of nucleotide substitution based on d_n cannot be compared directly; that is, the former is estimated by using all amino acid residues, whereas the latter is estimated by using only nonsynonymous nucleotide sites.

In comparing corresponding branch lengths among the three domains of mammalian ADH genes, CBD is frequently more conserved than CD1 (table 3). This result is consistent with the comparative analysis of amino acid sequences of mammalian ADHs by Yokoyama and Yokoyama (1987), who showed that the proportion of amino acid differences in CD1 and CD2 of human αADH is larger than that in CBD. Thus, for mammalian ADH genes, it appears that CBD is functionally more constrained than CD1. In this sense, HS4 is exceptional.

Branch lengths from C to various plant ADH genes is again variable (see fig. 1B); that is, branch D–HV2 is significantly longer than D–ZM1, D–HV1, and D–OS1. Similarly, HV2 is evolving significantly faster than HV3, TA, and PS (fig. 1B). Among the three functional domains, the branch lengths for CBD tend to be longer than those for CD1 (e.g., see table 3, branches C–AT, C–PS, D–OS2, and D–ZM1). Thus, in contrast to the situation in mammals, the CD1 region seems to be the most functionally constrained among the three domains in plant ADH genes.

Discussion

We have seen that in mammalian ADH genes more nonsynonymous substitutions have accumulated in CD1 than in CBD, whereas plant ADH genes show fewer substitutions in CD1 than in CBD and CD2. This can be studied more directly by using rates of nucleotide substitution.

To estimate such rates, we must know divergence times among plant and animal lineages. The time of the mammalian radiation is ~75 Myr ago (Mya). Monocots and dicots diverged from each other 100–200 Mya (Muller 1981; Shinozaki et al. 1983; Friis et al. 1987; Wolfe et al. 1989).

For mammals, separate estimates of evolutionary rates were obtained for primates (HS1, HS2, HS3, and PP) and murine (RN and MM) genes, by assuming that branch point B is 75 Mya (table 4). As expected, rates of nucleotide substitution in primate and murine ADH genes, particularly at nonsynonymous sites, tend to be lower within CBD than within CD1 (table 4). Similarly, CD2 is evolving more slowly than CD1 and CBD of murine genes and CD2 of primate genes (table 4).

For the entire coding sequence, the evolutionary rate for all nucleotide substitution

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1012357 by guest on 20 August 2022

Table 3
Branch Lengths Measured by d_n Values $\times 10^2$

Branch	Domain			All Domains
	CD1 (134 codons)	CBD (105 codons)	CD2 (42 codons)	(281 codons)
Mammals:				
A–HS4	15.3 ± 2.1	17.8 ± 2.6	14.1 ± 3.6	16.0 ± 1.5
A–HS2	15.0 ± 2.1	13.7 ± 2.2	13.7 ± 3.5	14.2 ± 1.4
B–HS1	7.8 ± 1.4[a]	3.4 ± 1.0[a]	4.9 ± 2.0	5.7 ± 0.8
B–HS2	5.0 ± 1.1	3.4 ± 1.0	4.3 ± 1.9	4.1 ± 0.7
B–MM	7.3 ± 1.4[b]	3.3 ± 1.0[b]	4.0 ± 1.8	5.3 ± 0.8
B–RN	7.1 ± 1.4	4.0 ± 1.1	2.0 ± 1.3	5.7 ± 0.8
Plants:				
C–AT	7.1 ± 1.4	9.3 ± 1.8	4.9 ± 2.0	7.6 ± 1.0
C–PS	4.8 ± 1.1	8.6 ± 1.7	10.8 ± 3.1	6.0 ± 0.9
D–HV2	5.3 ± 1.2	5.1 ± 1.3	11.0 ± 3.1	7.4 ± 0.9
D–OS2	4.0 ± 1.0	7.6 ± 1.6	4.7 ± 2.0	5.4 ± 0.8
D–ZM1	1.8 ± 0.7	4.5 ± 1.2	5.9 ± 2.2	3.4 ± 0.6

NOTE.—A, B, C, and D are as in table 2. Branch lengths between CD1 and CBD and between CD2 and CBD were compared separately.
[a] Difference in branch length is significant at the 5% level.
[b] Difference in branch lengths is significant at the 5% level.

is 1.5×10^{-9}/site/year for primate ADH and murine ADH genes (table 4). Furthermore, the evolutionary rates at synonymous sites are 5×10^{-9}/site/year and 4×10^{-9}/site/year for primates and murines, respectively. Thus, the evolution of ADH genes differs from a general pattern of nucleotide substitutions detected by Wu and Li (1985), who showed that synonymous nucleotide substitution is higher in murines than in humans.

If we assume that the monocot-dicot divergence occurred 200 Mya (Wolfe et al. 1989), then the rate of nucleotide substitution for the entire coding region of plant ADH genes is about half that for the entire coding region of mammalian genes. For example, the evolutionary rate is 0.8×10^{-9}/site/year for monocots (ZM1, HV1, and OS1) and dicots (AT and PS) (table 4). The two groups of genes show similar patterns of nucleotide substitutions; that is, the evolutionary rates at CD1, CBD, and CD2 for these genes are similar, and the average rate for the entire gene is 0.8×10^{-9}/site/year, 0.3×10^{-9}/site/year, and 3.0×10^{-9}/site/year for all codon positions, nonsynonymous sites, and synonymous sites, respectively. The duplicated genes HV2, HV3, TA, OS2, and ZM2 are evolving at a rate somewhat higher than these rates (table 4). The pattern of nucleotide substitution for these genes is characterized by the higher rate of synonymous substitution in CBD than in CD1 and CD2.

If we assume that monocots and dicots diverged 100 Mya, then the evolutionary rates for the coding regions of the mammalian and plant ADH genes are similar, both being $\sim 1.5 \times 10^{-9}$/site/year. The divergence times estimated by using molecular data are associated with large standard errors (e.g., see Wolfe et al. 1989), and therefore the possibility of equal evolutionary rates among the mammalian and plant ADH genes cannot be rejected.

In any case, it is clear that differential patterns of nucleotide substitutions exist among the three functional domains of different organisms. The variable evolutionary

Downloaded from https://academic.oup.com/mbe/article/7/2/143/... by guest on 20 August 2022

Table 4
Rates of Nucleotide Substitution/Site/Year × 10^9

	Domain			
Gene/Type of Sites	CD1	CBD	CD2	Combined
Primate:				
All	1.6 ± 0.2	1.4 ± 0.2	1.5 ± 0.4	1.5 ± 0.1
Nonsynonymous	0.8 ± 0.2	0.5 ± 0.1	0.5 ± 0.2	0.6 ± 0.1
Synonymous	5.0 ± 0.9	4.6 ± 0.9	5.9 ± 1.8	5.0 ± 0.6
Murine:				
All	1.6 ± 0.2	1.4 ± 0.2	0.8 ± 0.3	1.4 ± 0.1
Nonsynonymous	1.0 ± 0.2	0.6 ± 0.2	0.4 ± 0.2	0.7 ± 0.1
Synonymous	4.2 ± 0.8	4.2 ± 0.9	2.3 ± 1.0	4.1 ± 0.5
Monocot:				
a:				
All	0.8 ± 0.1	0.7 ± 0.1	0.8 ± 0.2	0.8 ± 0.1
Nonsynonymous	0.2 ± 0.1	0.2 ± 0.1	0.4 ± 0.1	0.3 ± 0.1
Synonymous	3.2 ± 0.5	2.8 ± 0.5	2.3 ± 0.7	2.9 ± 0.3
Monocot:				
b:				
All	0.6 ± 0.1	1.1 ± 0.1	1.0 ± 0.2	1.0 ± 0.1
Nonsynonymous	0.3 ± 0.1	0.3 ± 0.1	0.6 ± 0.2	0.4 ± 0.1
Synonymous	1.4 ± 0.3	5.4 ± 1.0	2.4 ± 0.7	4.0 ± 0.4
Dicot:				
All	0.7 ± 0.1	0.8 ± 0.1	0.8 ± 0.2	0.8 ± 0.1
Nonsynonymous	0.3 ± 0.1	0.4 ± 0.1	0.4 ± 0.1	0.3 ± 0.1
Synonymous	2.8 ± 0.5	2.1 ± 0.4	2.7 ± 0.8	2.8 ± 0.3

Note.—Branches B–HS1, B–HS2, B–HS3, and B–PP (primate); B–MM and B–RR (murine); C–OS1, C–ZM1, and C–HV1 (monocot a); C–HV3, C–TA, C–HV2, C–OS2, and C–ZM2 (monocot b); and C–AT and C–PS (dicot) in the tree of fig. 1 were considered. Divergence times of B and C were taken as 75 Mya and 200 Mya, respectively.

rates might have contributed to the functional adaptation of each ADH gene to its tissue-specific and developmental stage–specific expressions. However, such biological implications of the variable rates of nucleotide substitutions remain to be clarified.

Finally, when human ADH is used as a reference, amino acid residues 46 (Cys), 67 (His), and 174 (Cys)—which provide ligands to one catalytic zinc atom—and amino acid residues 97 (Cys), 100 (Cys), 103 (Cys), and 111 (Cys)—which provide ligands to another catalytic zinc atom—are identical for all mammalian and plant ADH genes.

Acknowledgments

Comments by Dr. M. Nei and anonymous reviewers were greatly appreciated. This work was supported by National Institutes of Health grants to S.Y. and R.Y., by a National Science Foundation grant to S.Y., and by United States Department of Agriculture grant 86-FSTY-9-0190 to D.E.H. and C.S.K. This work was also partially supported by grant DMB890002N and utilized the CRAY X-MP/48 system at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign.

Downloaded from https://academic.oup.com/mbe/article/7/2/141/1023575 by guest on 20 August 2022

LITERATURE CITED

Benyajati, C., A. R. Place, D. A. Powers, and W. Sofer. 1981. Alcohol dehydrogenase gene of *Drosophila melanogaster:* relationship of intervening sequences to functional domains in the protein. Proc. Natl. Acad. Sci. USA **78**:2717–2721.

Bühler, R., J. Hempel, R. Kaiser, C. de Zelenski, J.-C. von Wartburg, and H. Jörnvall. 1984. Human liver alcohol dehydrogenase. II. The primary structure of the γ_1 protein chain. Eur. J. Biochem. **145**:447–453.

Chang, C., and E. M. Meyerowitz. 1986. Molecular cloning and DNA sequence of the *Arabidopsis thaliana* alcohol dehydrogenase gene. Proc. Natl. Acad. Sci. USA **83**:1408–1412.

Crabb, D. W., and H. J. Edenberg. 1986. Complete amino acid sequence of rat liver alcohol dehydrogenase deduced from the cDNA sequence. Gene **48**:287–291.

Dennis, E. S., W. L. Gerlach, A. J. Pryor, J. L. Bennetzen, A. Inglis, D. Llewelyn, M. M. Sachs, R. J. Ferl, and W. T. Peacock. 1984. Molecular basis of the alcohol dehydrogenase (Adh1) gene of maize. Nucleic Acids Res. **12**:3983–4000.

Dennis, E. S., M. M. Sachs, W. L. Gerlach, E. J. Finnegan, and W. I. Peacock. 1985. Molecular analysis of the alcohol dehydrogenase 2 (Adh2) gene of maize. Nucleic Acids Res. **13**:727–743.

Edenberg, H. J., K. Zhang, K. Fong, W. F. Bosron, and T.-K. Li. 1985. Cloning and sequencing of cDNA encoding the complete mouse liver alcohol dehydrogenase. Proc. Natl. Acad. Sci. USA **82**:2262–2266.

Ehrig, T., J.-P. von Wartburg, and B. Wermuth. 1988. cDNA sequence of the β_2-subunit of human liver alcohol dehydrogenase. FEBS Lett. **234**:53–55.

Eklund, H., C.-I. Brändén, and H. Jörnvall. 1976*a*. Structure comparisons of mammalian, yeast and bacillar alcohol dehydrogenase. J. Mol. Biol. **102**:61–73.

Eklund, H., B. Nordström, E. Zeppezauer, G. Söderlund, I. Ohlsson, T. Boine, B.-O. Söderberg, O. Tapia, C.-I. Brändén, and Å. Åkeson. 1976*b*. Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution. J. Mol. Biol. **102**:27–59.

Freeling, M., and D. Bennett. 1985. Maize *Adh1*. Annu. Rev. Genet. **19**:297–323.

Friis, E., W. Chalconer, and P. Crane. 1987. Introduction to angiosperms. Pp. 1–15 *in* Friis, E., W. Chalconer, and P. Crane, eds. The origins of angiosperms and their biological consequences. Cambridge University Press, Cambridge.

Gojobori, T., K. Ishii, and M. Nei. 1982. Estimation of average number of nucleotide substitution varies with nucleotide. J. Mol. Evol. **18**:414–423.

Good, A. G., L. E. Pelcher, and W. L. Crosby. 1988. Nucleotide sequence of a complete barley alcohol dehydrogenase 1 cDNA. Nucleic Acids Res. **16**:7182.

Harberd, N., and K. Edwards. 1983. Further studies on the alcohol dehydrogenases in barley: evidence for a third alcohol dehydrogenase locus and the data on the effect of an *alcohol dehydrogenase-1 null* mutation in homozygous and in heterozygous condition. Genet. Res. **41**:109–116.

Hempel, J., R. Bühler, R. Kaiser, B. Holmquist, C. de Zalenski, J.-P. von Wartburg, B. Vallee, and H. Jörnvall. 1984. Human liver alcohol dehydrogenase. I. The primary structure of the $\beta_1\beta_1$ isoenzyme. Eur. J. Biochem. **145**:437–445.

Höög, J.-O., L.-O. Hedén, K. Larsson, H. Jörnvall, and H. von Bahr-Lindström. 1986. The γ_1 and γ_2 subunits of human liver alcohol dehydrogenase. Eur. J. Biochem. **159**:215–218.

Höög, J.-O., H. von Bahr-Lindström, L.-O. Hedén, B. Holmquist, K. Larsson, J. Hempel, B. L. Vallee, and H. Jörnvall. 1987. Structure of the class II enzyme of human liver alcohol dehydrogenase: combined cDNA and protein sequence determination of the π subunit. Biochemistry **26**:1926–1932.

Ikuta, E., S. Szeto, and A. Yoshida. 1986. Three human alcohol dehydrogenase subunits:

Downloaded from https://academic.oup.com/mbe/article/7/143/1012377 by guest on 20 August 2022

cDNA structure and molecular and evolutionary divergence. Proc. Natl. Acad. Sci. USA **83:** 634–638.

JÖRNVALL, H. 1985. Alcohol dehydrogenase, aldehyde dehydrogenase, and related enzymes. Alcohol **2:**61–66.

JÖRNVALL, H., H. BAHR-LINDSTRÖM, K.-D. JANY, W. ULMER, and M. FROSCHLE. 1984. Extended superfamily of short alcohol-polyol-sugar dehydrogenases: structural similarities between glucose and ribitol dehydrogenases. FEBS Lett. **165:**190–196.

JÖRNVALL, H., M. PERSSON, and J. JEFFREY. 1981. Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. Proc. Natl. Acad. Sci. USA **78:**4226–4230.

———. 1987. Characteristics of alcohol/polyol dehydrogenases: the zinc-containing long-chain alcohol dehydrogenases. Eur. J. Biochem. **167:**195–201.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. MUNRO, ed. Mammalian protein metabolism. III. Academic Press, New York.

JULIA, P., X. PARÉS, and H. JÖRNVALL. 1988. Rat liver alcohol dehydrogenase of class III. Eur. J. Biochem. **172:**73–83.

KIMURA, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. USA **78:**454–458.

LLEWELLYN, D. J., E. J. FINNEGAN, J. G. ELLIS, E. S. DENNIS, and W. J. PEACOCK. 1987. Structure and expression of an alcohol dehydrogenase 1 gene from *Pisum sativum* (c. v. "Greenfast"). J. Mol. Biol. **195:**115–123.

MATSUO, Y., R. YOKOYAMA, and S. YOKOYAMA. 1989. Human alcohol dehydrogenases β1 and β2 can be specified by a single nucleotide difference. Eur. J. Biochem. **183:**317–320.

MATSUO, Y., and S. YOKOYAMA. 1989. Molecular structure of alcohol dehydrogenase 1 gene. FEBS Lett. **243:**57–60.

MITCHELL, L. E., E. S. DENNIS, and W. J. PEACOCK. 1989. Molecular analysis of an alcohol dehydrogenase (Adh) gene from chromosome 1 of wheat. Genome **32:**349–358.

MIYATA, T., and T. YASUNAGA. 1980. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitution from homologous nucleotide sequences and its application. J. Mol. Evol. **16:**23–26.

MULLER, J. 1981. Fossil pollen records of extant angiosperms. Bot. Rev. **47:**1–142.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

SACHS, M., E. DENNIS, W. GERLACH, and W. PEACOCK. 1986. Two alleles of maize *Alcohol dehydrogenase 1* have 3′ structural and poly(A) addition polymorphisms. Genetics **116:**449–467.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.

SHINOZAKI, K., C. YAMADA, N. TAKAHATA, and M. SUGIURA. 1983. Molecular cloning and sequence analysis of the cyanobacterial gene for the large subunit of ribulose-1, 5-biphosphate carboxylase/oxydase. Proc. Natl. Acad. Sci. USA **80:**4050–4054.

SMITH, M. 1986. Genetics of human alcohol and aldehyde dehydrogenases. Adv. Hum. Genet. **14:**249–290.

SOKAL, R. R., and C. D. MICHENER. 1958. A statistical method for evaluating systematic relationships. Univ. Kans. Sci. Bull. **28:**1409–1438.

TAJIMA, F., and M. NEI. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Biol. Evol. **1:**269–285.

TAKAHATA, N., and M. KIMURA. 1981. A method of evolutionary base substitutions and its application with special reference to rapid change of pseudogenes. Genetics **98:**641–657.

TREZISE, A. E. O., E. A. GODFREY, R. S. HOLMES, and I. R. BEACHAM. 1989. Cloning and sequencing of cDNA encoding baboon liver alcohol dehydrogenase: evidence for a common ancestral lineage with the human alcohol dehydrogenase β subunit and for class I ADH gene duplications predating primate radiation. Proc. Natl. Acad. Sci. USA **86:**5454–5458.

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1023557 by guest on 20 August 2022

TRICK, M., E. S. DENNIS, K. J. R. EDWARDS, and W. J. PEACOCK. 1988. Molecular analysis of the alcohol dehydrogenase gene family of barley. Plant Mol. Biol. **11**:147–160.

WOLFE, K. H., M. GOUY, Y.-W. YANG, P. M. SHARP, and W.-H. LI. 1989. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proc. Natl. Acad. Sci. USA **86**:6201–6205.

WU, C.-I., and W.-H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA **82**:1741–1745.

XIE, Y., and R. WU. 1989. Rice alcohol dehydrogenase genes: anaerobic induction, organ specific expression and characterization of cDNA clones. Plant Mol. Biol. **13**:53–68.

YOKOYAMA, S., and R. YOKOYAMA. 1987. Molecular evolution of mammalian class I alcohol dehydrogenase. Mol. Biol. Evol. **4**:504–513.

YOKOYAMA, S., R. YOKOYAMA, and P. ROTWEIN. 1987. 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. Jpn. J. Genet. **62**:241–256.

MASATOSHI NEI, reviewing editor

Received April 5, 1989; revision received May 30, 1989

Accepted October 24, 1989

Downloaded from https://academic.oup.com/mbe/article/7/2/143/1012357 by guest on 20 August 2022