Cytosolic Isocitrate Dehydrogenase in Humans, Mice, and Voles and Phylogenetic Analysis of the Enzyme Family

Anton Nekrutenko, David M. Hillis,† John C. Patton,* Robert D. Bradley,*and Robert J. Baker**

Department of Biological Sciences, Texas Tech University; and †Department of Zoology and Institute of Cellular and Molecular Biology, University of Texas at Austin

In this study, we report cDNA sequences of the cytosolic NADP-dependent isocitrate dehydrogenase for humans, mice, and two species of voles (*Microtus mexicanus* and *Microtus ochrogaster*). Inferred amino acid sequences from these taxa display a high level of amino acid sequence conservation, comparable to that of myosin β heavy chain, and share known structural features. A *Caenorhabditis elegans* enzyme that was previously identified as a protein similar to isocitrate dehydrogenase is most likely the NADP-dependent cytosolic isocitrate dehydrogenase enzyme equivalent, based on amino acid similarity to mammalian enzymes and phylogenetic analysis. We also suggest that NADP-dependent isocitrate dehydrogenases characterized from alfalfa, soybean, and eucalyptus are most likely cytosolic enzymes. The phylogenetic tree of various isocitrate dehydrogenases from eukaryotic sources revealed that independent gene duplications may have given rise to the cytosolic and mitochondrial forms of NADP-dependent isocitrate dehydrogenase in animals and fungi. There appears to be no statistical support for a hypothesis that the mitochondrial and cytosolic forms of the enzyme are orthologous in these groups. A possible scenario of the evolution of NADP-dependent isocitrate dehydrogenases is proposed.

Introduction

Eukaryotic cells express three different isocitrate dehydrogenases that catalyze decarboxylation of isocitrate into α-ketoglutarate. Genes that encode these three enzymes are located in the nuclear genome, although their protein products function in the cytoplasm (Jennings et al. 1994), in mitochondria (Zhao and McAlister-Henn 1996), and in plastids (Chen et al. 1989). These enzymes utilize either NAD or NADP as cofactors. The NAD-dependent mitochondrial enzyme catalyzes a key step in the tricarboxylic acid cycle, whereas the physiological roles of two NADP-dependent enzymes (mitochondrial ICDH and cytosolic ICDH, depending on the localization of the functional enzyme) are not clearly understood. The NAD-dependent mitochondrial enzyme (EC 1.1.1.41) is a heterooctamer $\alpha_4\beta_2\gamma_2$ ($\alpha_4\beta_4$ in yeast), and each subunit is encoded by a separate nuclear gene. Both NADP-dependent enzymes (EC 1.1.1.42) are homodimers that are also encoded in the nuclear genome (Ramachandran and Colman 1980; Keys and McAlister-Henn 1990). Gene disruption studies in yeast have shown that NADP-dependent enzymes cannot compensate for the function of the NAD-specific isocitrate dehydrogenase (Haselbeck and McAlister-Henn 1993). There is evidence that the cytosolic enzyme participates in the production of NADPH (Winkler, Desantis, and Solomon 1986) and in the biosynthesis of fatty acids (Belfiore and Iannello 1995).

In contrast to eukaryotes, *Escherichia coli* cells contain a single isocitrate dehydrogenase enzyme; it is an NADP-dependent homodimer involved in the tricarboxylic acid cycle and is regulated by the phosphorylation of the active center (Hurley et al. 1990). It is the only isocitrate dehydrogenase for which the three-dimensional structure has been determined by X-ray crystallography (Hurley et al. 1991). Sharing only slight amino acid similarity with the eukaryotic NADP-dependent enzymes, it is nonetheless capable of restoring the function to the cytosolic or mitochondrial ICDH in yeast cells carrying disruption of either gene. At the same time, it is not competent for restoration of the respiratory function to yeast cells with a disrupted gene for the NAD-dependent mitochondrial enzyme (Zhao and McAlister-Henn 1996). Recently, a novel isocitrate dehydrogenase was characterized from an α-group proteobacterium, *Sphingomonas yanoikuyae* (Wang and Lau 1996). This polypeptide has an unexpectedly high degree of similarity to eukaryotic enzymes (table 1): the amino acid identity between the *E. coli* isocitrate dehydrogenase and mammalian NADP-dependent isocitrate dehydrogenases averages 14% (Haselbeck, Colman, and McAlister-Henn 1992), whereas the identity between the *S. yanoikuyae* enzyme and the human NADP-dependent cytosolic enzyme is 62%.

ICDHs, and particularly the NADP-dependent cytosolic enzymes, are the focus of this study. Inferred amino acid sequences of the cytosolic enzyme were previously known only for the laboratory rat (*Rattus norvegicus*; Jennings et al. 1994), for yeast (*Saccharomyces cerevisiae*; Loftus et al. 1994), and for two species of plants, potato (*Solanum tuberosum*; Fieuw et al. 1995) and tobacco (*Nicotiana tabacum*; Galvez et al. 1996). The mitochondrial NADP-dependent isocitrate dehydrogenase was characterized for humans (*Homo sapiens*; Huh et al., unpublished data from ENTREZ database; see table 1), mice (*Mus musculus*; Yang et al. 1996), pigs (*Sus scrofa*; Haselbeck, Colman, and McAlister-Henn 1992), cows (*Bos taurus*; Huh et al. 1993), and yeast (*S. cerevisiae*; Haselbeck and McAlister-Henn 1991). The NADP-dependent isocitrate dehydrogenase has also been characterized for eucalyptus (*Eucalyptus globulus*; Boiffin et al., unpublished data from ENTREZ

Key words: NADP-dependent cytosolic isocitrate dehydrogenase, ICDH, human, mouse, *Microtus.*

Address for correspondence and reprints: Anton Nekrutenko, Department of Biological Sciences,Texas Tech University, Lubbock, Texas 79409. E-mail: anton@ttu.edu.

Mol. Biol. Evol. 15(12):1674–1684. 1998
© 1998 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Table 1
Sequences of NADP-Dependent Isocitrate Dehydrogenase Available from the ENTREZ Database

Enzyme Source	Accession Number	Author(s)
Rat cytosolic	L35317	Jennings et al. (1994)
Yeast cytosolic	P41939	Loftus et al. (1994)
Potato cytosolic	P50217	Fieuw et al. (1995)
Tobacco cytosolic	P50218	Galvez et al. (1996)
Human mitochondrial	P48735	Huh et al. (unpublished data)
Mouse mitochondrial	P54071	Yang et al. (1996)
Pig mitochondrial	P33198	Haselbeck, Colman, and McAlister-Henn (1992)
Cow mitochondrial	Q04467	Huh et al. (1993)
Yeast mitochondrial	P21954	Haselbeck and McAlister-Henn (1991)
Nematode	Z68343	Wilson et al. (1994)
Alfalfa	M93672	Shorrosh and Dixon (1992)
Eucalyptus	U80915	Boiffin et al. (unpublished data)
Soybean	Q06197	Udvardi, McDermott, and Kahn (1993)
Yeast ORF YNL009W	Z71285	Andre et al. (unpublished data)
α-proteobacterium	P50215	Wang and Lau (1996)

database; see table 1), soybean (*Glycine max*; Udvardi, McDermott, and Kahn 1993), and alfalfa (*Medicago sativa*; Shorrosh and Dixon 1992). However, no subcellular localization has been proposed for the enzymes from these plants.

In this paper, we report and characterize cDNA sequences that encode the NADP-dependent cytosolic isocitrate dehydrogenase in humans (*H. sapiens*), mouse (*M. musculus*), and two species of voles (*Microtus mexicanus* and *Microtus ochrogaster*). We also summarize the available sequence data in a phylogenetic perspective.

Materials and Methods

Total RNA Isolation and Reverse Transcription–Polymerase Chain Reaction (RT-PCR)

We used reverse transcription–polymerase chain reaction (RT-PCR) for initial characterization of human NADP-dependent cytosolic isocitrate dehydrogenase cDNA and for isolation of the mouse and vole cDNAs. Total RNA was isolated from human placenta, mouse liver, and vole livers using the ToTally RNA extraction kit (Ambion, Inc.). For the first strand of cDNA synthesis, the M-MLV reverse transcriptase and Oligo-$(dT)_{15}$ primer (Promega Corp.) were used. First-strand synthesis was followed by PCR amplification with isocitrate dehydrogenase–specific primers using *Taq* DNA polymerase (Promega Corp.) and DNA Thermal Cycler (Perkin Elmer Corp.). The first pair of primers for that purpose was developed using a cDNA sequence of rat cytosolic isocitrate dehydrogenase (Jennings et al. 1994). Only cDNA corresponding to the cytosolic enzyme can be amplified using these primers. Additional primers were developed for amplification and/or sequencing of cDNA from different species. Sequences of primers and the reaction conditions can be obtained from World Wide Web site http://www.nsrl.ttu.edu/icdh.htm or from A.N. on request (anton@ttu.edu). PCR products were cloned using the pGEM-T vector (Promega Corp.) or were sequenced directly.

Screening of Human cDNA Library

In order to obtain the full-length cDNA of the human NADP-dependent cytosolic isocitrate dehydrogenase, we screened an adult liver cDNA library (Clontech Laboratories, Inc.) by PCR. The 5′-end of the cDNA was amplified using a primer complementary to the vector sequence and a cytosolic ICDH-specific primer. As a result, we obtained a set of PCR products of various sizes. The products were subcloned into the pGEM-T, and the clone with the largest insert was selected. Two-step PCR (Yan, Yang, and Parkinson 1995) was used to subclone the 3′ end using one vector-complementary and two isocitrate dehydrogenase–specific primers. The single band obtained was then subcloned using the pGEM-T vector system. Nested deletions were then generated for this clone using Exo III and S I nucleases (Promega Corp.) as described in manufacturer's protocol to allow sequencing of the entire 3′ region.

DNA Sequencing and Sequence Analysis

PCR products were sequenced using the Dye Terminator Cycle Sequencing Kit or the dRhodamine Dye Terminator Cycle Sequencing Kit (Perkin Elmer Corp.) and analyzed on an ABI Prism 310 automated sequencer (Perkin Elmer Corp.). A large number of PCR products from independent reactions were analyzed to detect possible amplification artifacts. All reported sequences were determined by sequencing in both directions. The Sequencher 3.0 software package (GeneCodes Corp.) was used to analyze and edit sequence data. Translation of nucleotide sequences and analysis of amino acid sequences (amino acid composition, *p*I value calculation, and prediction of secondary structure) was performed using the computer package MacVector 5.0 (Oxford Molecular Group). Nucleotide sequences were deposited in GenBank (human, AF020038; mouse, AF020039; *M. mexicanus*, AF048831; *M. ochrogaster*, AF048832). Alignment of amino acid sequences was conducted using the CLUSTAL W program (Thompson, Higgins, and Gibson 1994). Protein motifs were identified using the

PROSITE database (Bairoch et al. 1997) and the PSORT II program, developed by Nakai (1997).

Sequence Data from Other Sources

In addition to the sequences reported in this study, we used data available from the ENTREZ database. Descriptions, accession numbers, and authors of sequences used are listed in table 1.

Phylogenetic Analyses

Phylogenetic analyses were performed using test versions 4.0d61a–4.0d63 of PAUP* (Swofford 1998). Analyses of amino acid sequences were conducted using the PROTPARS method with weighted parsimony (Felsenstein 1993; see Swofford et al. 1996). The *S. yanoikuyae* (proteobacterium) sequence was selected as an outgroup for this analysis. Maximum-likelihood analyses of a more limited set of DNA sequences were based on the General Time-Reversible Model, with all six substitution types, the base frequencies, and the number of invariant sites estimated from the data (see Swofford et al. 1996). Best trees were found by stepwise addition of taxa followed by tree-bisection-reconnection branch-swapping (Swofford et al. 1996). Nonparametric bootstrapping analysis (Felsenstein 1985) was used to determine the level of support for each clade in the parsimony analyses. Alternative hypotheses were tested in the maximum-likelihood analyses by constraining the search to trees that fit the null hypothesis, and then comparing the log-likelihood values of the constrained tree with those of the optimal tree (Hillis, Mable, and Moritz 1996; Huelsenbeck, Hillis, and Nielsen 1996).

Results and Discussion

Cloning of the Cytosolic ICDH

We cloned cDNAs encoding cytosolic ICDH from humans, mice, and two species of voles (*M. arvalis* and *M. rossiaemeridionalis*). Our conclusion that we have isolated sequences encoding the cytosolic enzyme from these taxa is based on the high degree of identity between these sequences and the sequence of the rat enzyme characterized previously (Jennings et al. 1994; see below). Alignment of these nucleotide sequences, together with the cDNA sequence of the cytosolic ICDH from the laboratory rat, can be obtained at World Wide Web site http://www.nsrl.ttu.edu/icdh.htm or directly from A.N. The mouse and two vole cytosolic ICDH cDNAs that we sequenced contained only short portions of the 5′- and 3′-untranslated regions (UTRs), compared to the rat sequence. These rodent cDNAs, as well as the human cDNA, each have predicted coding regions of 1,245 nt in length. The degree of nucleotide identity among these coding regions ranges from a minimum of 90.12% between human and mouse sequences to a maximum of 99.03% between the two vole sequences.

The human ICDH cDNA that we cloned is 2,343 nt in length (249 nt of 5′-UTR, 1,245 nt of the coding region, and 849 nt of 3′-UTR), excluding the poly(A) tail present at the 3′ end. A single putative polyadenylation signal is found 603 nt downstream of the stop codon and 240 nt upstream of the polyadenylation site. Such a long distance between the polyadenylation signal and the polyadenylation site is much greater than the typical distance found in most mRNAs, which is about 15 nt on average (Wahle and Keller 1996). It is possible that either there is another, cryptic, polyadenylation signal or that this mRNA folds into a conformation that brings the polyadenylation signal and site into closer proximity.

To determine whether we cloned a full-length human cytosolic ICDH transcript, we compared it against the expressed sequence tag (EST) database (dbest) at the National Center for Biotechnology Information. This analysis revealed a large number of ESTs corresponding to different portions of our sequence, but none of them extended beyond the 5′ end of the human ICDH cDNA that we cloned. We also found two ESTs corresponding to the extreme 3′ end of our sequence (accession numbers AA666366 and AA679791). Both of these ESTs contained the polyadenylation signal, and separated from the polyadenylation site by 240 nt—the same pattern that is found in the human ICDH cDNA described here. Furthermore, we did not find any polyadenylated ESTs that contained poly(A) tails closer to the putative polyadenylation site. These data allow us to conclude that we cloned a full-length cDNA of the human cytosolic ICDH.

The coding regions of the cytosolic ICDH cDNAs of humans, mice, and voles encode a single subunit (ICDH is a homodimer) of the enzyme that is 414 amino acids long. The average calculated molecular weight for this protein is 46,671 Da (46,685 Da for the human, 46,656 Da for the mouse, 46,656 Da for *M. ochrogaster*, and 46,690 Da for *M. mexicanus*). All of these proteins contain a C-terminal tripeptide, alanine-lysine-leucine (fig. 1), which is the type 1 peroxisomal targeting sequence (McNew and Goodman 1996). The amino acid identities between the various ICDHs are given in table 2.

The isolation and characterization of mouse and human cytosolic ICDH cDNAs has an important outcome: we now have three taxa (humans, mice, and yeast) for which sequences of both cytosolic and mitochondrial ICDHs are available. This provides a basis for the phylogenetic analyses described below. The sequences from voles, on the other hand, further demonstrate the degree of nucleotide and amino acid conservation of the cytosolic ICDHs among mammalian species.

Search of the Yeast Genome Database for Additional ICDH Loci

The yeast genome database, recently released at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/Entrez/Genome/org.html), provides a unique opportunity to search the entire genome for additional members of the isocitrate dehydrogenase family. To do so, we performed a BLAST search (http://www.ncbi.nlm.nih.gov/BLAST) of that database using nucleotide sequences of coding regions corresponding to the yeast mitochondrial NADP-dependent (Haselbeck and McAlister-Henn 1991) and cytosolic NADP-dependent (Loftus et al. 1994) enzymes. This analysis re-

vealed three ICDH-like sequences in the yeast genome under the following accession numbers: U17246, a region of *S. cerevisiae* chromosome XII cosmid 9470 (Johnston et al. 1997); Z71285, representing *S. cerevisiae* chromosome XIV reading frame ORF YNL009W (Andre et al., unpublished data from the ENTREZ database); and Z74114, a reading frame ORF YDL066W from *S. cerevisiae* chromosome IV (Bloecker and Brandt, unpublished data from the ENTREZ database).

An open reading frame (ORF) within U17246 (positions 6755–7993) has 96% nucleotide identity to the coding region of the previously characterized gene for the yeast cytosolic enzyme (Loftus et al. 1994). It is designated the ''cytosolic form of NADP-dependent isocitrate dehydrogenase'' by authors of the sequence. The predicted protein product encoded by that sequence (PID: g577204) is 97% identical to the amino acid sequence of the yeast cytosolic enzyme. It appears, based on such identity, that the sequence published by Loftus et al. (1994) and the sequence g577204 represent the same locus. The sequence Z74114 contains an ORF (positions 439–1725) which is 99% identical at the DNA sequence level to the coding region of the yeast mitochondrial NADP-dependent enzyme gene characterized previously (Haselbeck and McAlister-Henn 1991). The predicted protein product of the reading frame (PID: g1431074) is identical to the sequence published by Haselbeck and McAlister-Henn (1991).

The ORF YNL009W (accession number Z71285) is noteworthy. It is only 66% identical to the published nucleotide sequence of the coding region of the yeast cytosolic ICDH and 62% identical to the sequence of the yeast mitochondrial ICDH coding region (table 2 and fig. 1). The amino acid sequence identity values are 62% compared to the cytosolic enzyme and 52% compared to the mitochondrial ICDH. The predicted protein product encoded by ORF YNL009W has unique structural properties: (1) it is a basic polypeptide (estimated $pI = 9.16$); (2) it does not have the mitochondrial signal peptide, as indicated by PSORT analysis (Nakai 1997); and (3) it appears to have the type 1 peroxisomal targeting signal (PTS1) at the C-terminus (fig. 1). These features are somewhat contrary to the general pattern observed for cytosolic and mitochondrial ICDHs (see *Discussion*) in that the protein encoded by ORF YNL009W combines properties characteristic of both mitochondrial and cytosolic forms of yeast ICDH. Yeast mutants carrying disrupted genes for both cytosolic and mitochondrial NADP-dependent isocitrate dehydrogenases do not show measurable ICDH activity, and therefore any product of ORF YNL009W does not appear to be a functional ICDH (McAlister-Henn and Small 1997). If this sequence is included in the phylogenetic analysis (fig. 2), it clusters together with the cytosolic ICDH from yeast. Given the apparent lack of detectable ICDH activity from this product, ORF YNL009W may represent either a nonfunctional gene duplication or a gene duplication that has taken on another (as yet unknown) function.

Conserved Regions in Eukaryotic ICDHs

Hurley et al. (1991) identified amino acid residues of the *E. coli* isocitrate dehydrogenase involved in binding of the isocitrate-Mg^{2+} complex. Using these data, Jennings et al. (1994) compared the amino acid sequence of the rat ICDH with the sequence of the *E. coli* enzyme and identified seven potentially homologous binding sites. Our comparison indicates that these seven amino acid sites are identical for all ICDH sequences available to date (fig. 1). The pairwise amino acid identity matrix for all known eukaryotic ICDHs and the proteobacterial (*S. yanoikuyae*) ICDH is shown in table 2. Overall, identical amino acid sites of the compared polypeptides within the species examined are organized in blocks rather than randomly distributed. These regions may represent catalytically active portions of the enzyme, may participate in the assemblage of subunits, and/or may participate in polypeptide folding. However, there is no obvious correlation between the positions of the amino acids possibly involved in binding of the isocitrate-Mg^{2+} complex discussed earlier and the positions of these conserved blocks (fig. 1). The enzyme from a very distant organism, the proteobacterium *S. yanoikuyae,* has the same pattern of conservative amino acid sites. For example, there are 157 identical amino acid positions among all eukaryotic ICDH proteins, and 152 of these also are also identical in the *S. yanoikuyae* polypeptide. Of the remaining 5 amino acids, which are different in the *S. yanoikuyae* enzyme compared with eukaryotic polypeptides, 2 are conservative substitutions.

When compared with other proteins, cytosolic ICDH displays a high degree of amino acid conservation. We performed the simple analysis described below in order to illustrate the conservatism of this polypeptide. Li (1997) published rates of nonsynonymous substitutions for various mammalian protein-coding genes, including data for the myosin β heavy chain. The myosin β heavy chain is a very conservative protein based on the rate of nonsynonymous substitutions in the coding region of its gene. When amino acid sequences of human and rat myosins (Jaenicke et al. 1990; Kraft et al. 1989) are compared, there are 56 variable amino acid sites (3%) in the 1,936 amino acids representing the entire molecule. A comparison of amino acid sequences from human and rat cytosolic isocitrate dehydrogenase depicts 18 variable sites (4.3%) from a total of 414 sites. These values (3% of variable sites in the myosin molecule and 4.3% in the cytosolic isocitrate dehydrogenase) are comparable, proving that the cytosolic ICDH is a very conserved protein. This fact emphasizes the physiological importance of the cytosolic isocitrate dehydrogenase in eukaryotic cells.

Protein Sequence Motifs

To summarize the information concerning various protein sequence motifs found in ICDHs, we compared inferred amino acid sequences reported in this study, as well as amino acid sequences procured from the ENTREZ database (table 1), against the PROSITE subsequence database. These analyses revealed several interesting features. First, as expected, all NADP-dependent

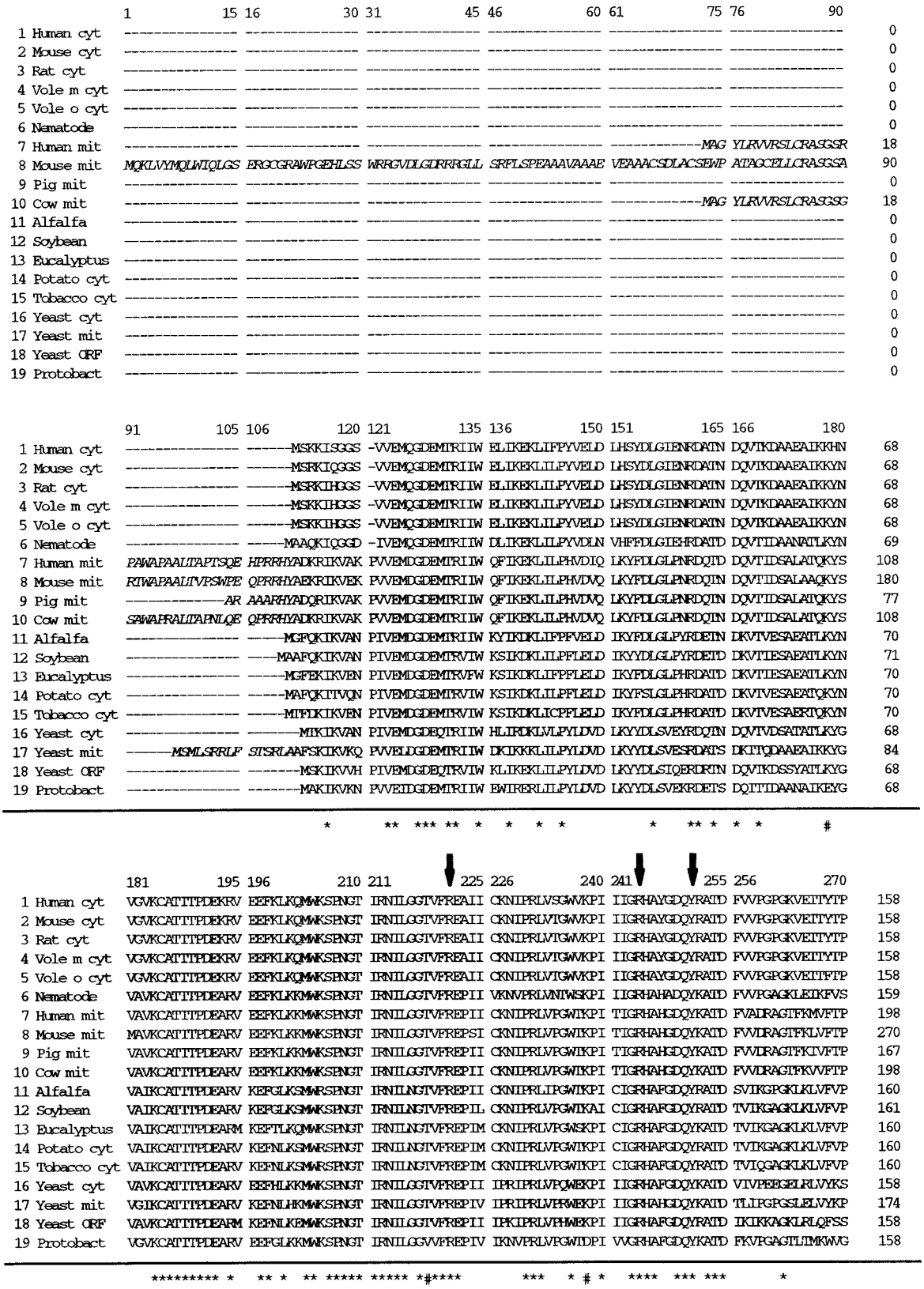

FIG. 1.—Alignment of amino acid sequences of NADP-dependent isocitrate dehydrogenases presented in this study and available from GenBank (table 1). The numbers at the top of each alignment block denote the scale of the figure. The number at the end of each line indicates the position of the last amino acid within that line. Gaps in the alignment are indicated by dashes. Signal peptides of the mitochondrial enzyme are italicized. The isocitrate dehydrogenase and isopropylmalate dehydrogenase signature sequence is indicated by a vertical rectangle, and peroxisomal targeting sequences are indicated by horizontal rectangles. Amino acids possibly involved in isocitrate binding are highlighted by arrows. The asterisks at the bottom of each alignment block signify that all sequences have identical residues at that position. Abbreviations: #, all eukaryotic sequences (1—18) have identical residues; cyt, the cytosolic enzyme; mit, the mitochondrial enzyme; Vole m, *M. mexicanus*; Vole o, *M. ochrogaster*; Nematode, *C. elegans*; Yeast ORF, yeast open reading frame YNL009W; Protobact., *S. yanoikuyae.*

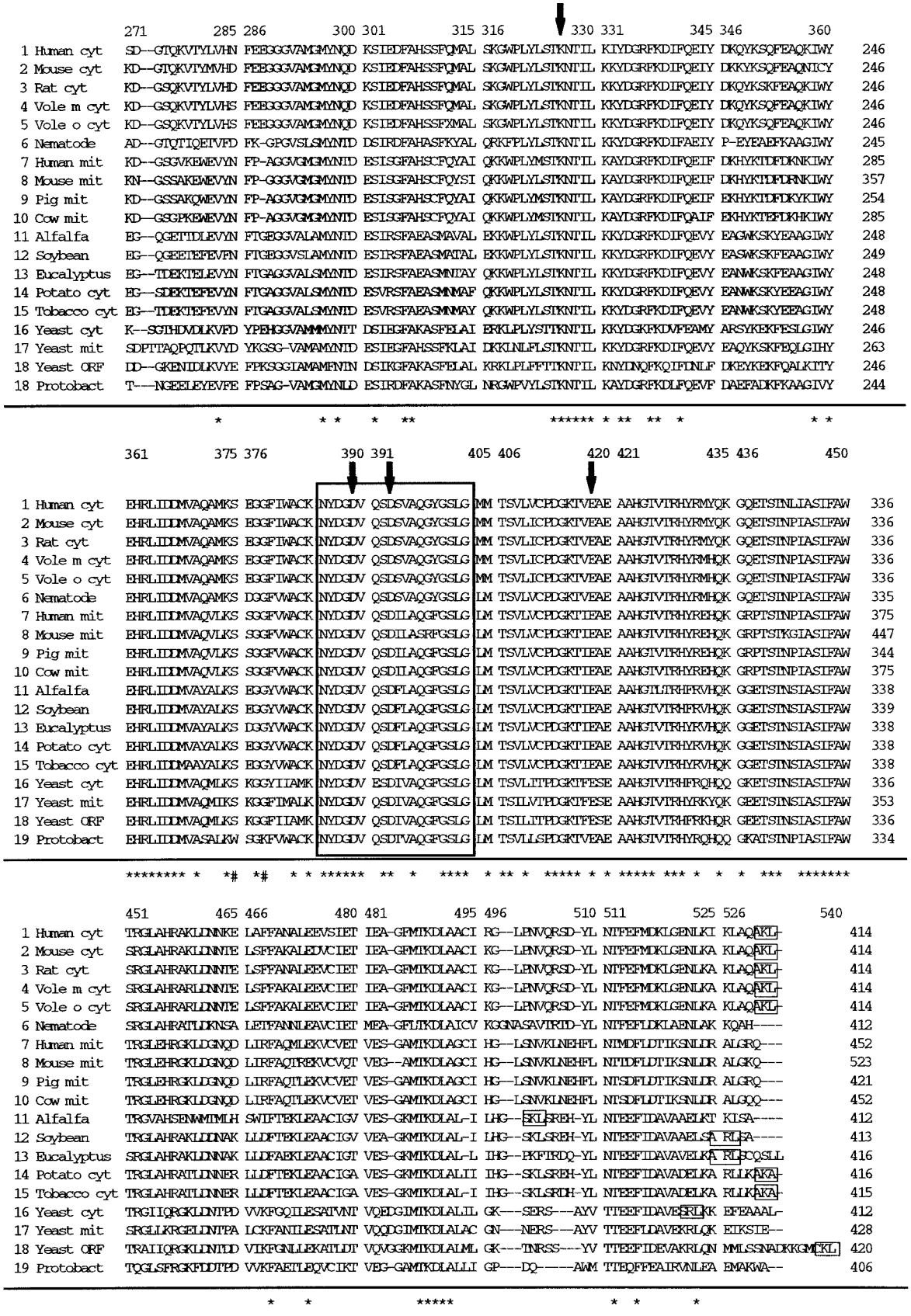

FIG. 1 *(Continued)*

Table 2
Deduced Amino Acid Sequence Identities Among NADP-Dependent Isocitrate Dehydrogenases

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Human cyt...	—																		
2. Mouse cyt ...	94.7	—																	
3. Vole m cyt...	95.4	97.3	—																
4. Vole o cyt ...	95.2	97.1	99.8	—															
5. Rat cyt.....	95.7	97.8	98.6	98.3	—														
6. Nematode....	75.1	76.5	76.3	76.7	76.5	—													
7. Human mit ..	68.9	68.6	69.6	69.8	69.3	68.3	—												
8. Mouse mit...	67.6	67.1	68.0	68.2	67.8	66.0	90.2	—											
9. Pig mit.....	69.1	69.1	69.8	70.0	69.8	69.0	96.0	93.1	—										
10. Cow mit.....	69.1	68.9	69.8	70.0	69.6	68.3	96.0	90.2	96.7	—									
11. Alfalfa	64.1	63.6	64.1	64.5	64.1	65.0	65.6	63.8	66.3	65.9	—								
12. Soybean.....	65.3	65.5	65.8	66.2	66.1	67.7	68.1	66.1	68.1	67.9	89.8	—							
13. Tobacco cyt...	64.8	64.8	65.3	65.7	65.3	66.3	67.6	65.9	67.4	67.4	86.2	89.1	—						
14. Potato cyt ...	64.9	65.1	65.6	66.0	65.6	67.1	68.0	65.9	68.0	67.7	86.9	90.3	97.1	—					
15. Eucalyptus...	65.8	66.0	66.3	66.7	66.5	68.9	68.6	66.6	68.4	68.4	84.7	89.1	90.8	90.1	—				
16. Yeast cyt	61.2	62.2	62.0	61.9	62.4	62.0	63.2	61.9	64.0	64.2	64.0	65.0	61.8	62.0	63.0	—			
17. Yeast mit	65.0	65.3	65.0	65.0	66.0	63.3	61.0	60.0	62.5	61.5	63.3	65.1	64.5	64.7	66.2	73.4	—		
18. Yeast ORF...	59.7	59.4	59.7	60.0	60.1	60.4	61.9	60.6	62.2	62.2	62.2	63.6	61.0	61.5	62.5	75.5	70.0	—	—
19. Proteobact ...	62.5	63.5	63.5	63.7	63.8	64.4	66.0	64.9	66.0	66.0	61.8	65.6	62.5	62.6	64.0	64.6	63.9	61.6	—

NOTE.—Abbreviations: cyt, cytosolic; mit, mitochondrial; Vole o, *Microtus ochrogaster*; Vole m, *Microtus mexicanus*; Nematode, *Caenorhabditis elegans*; Yeast ORF, open reading frame YNL009W; Proteobact; *Sphingomonas yamoikuyae.*

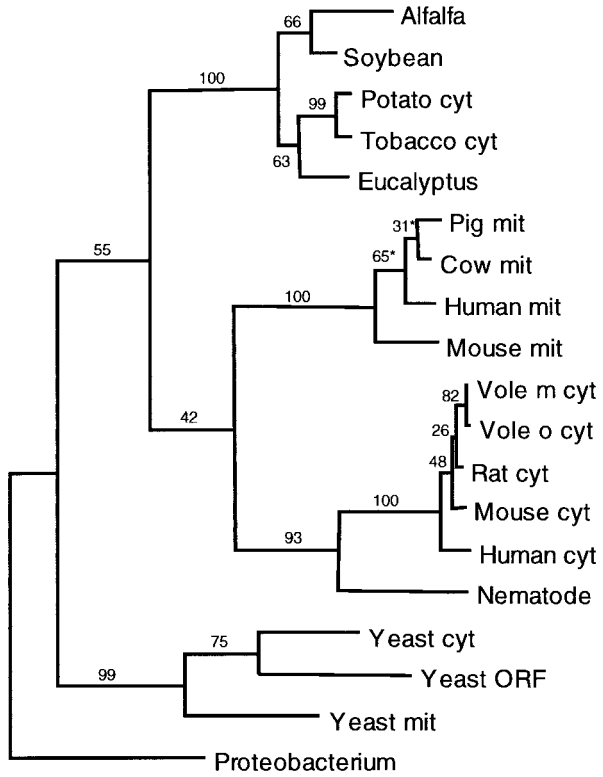

FIG. 2.—Estimated phylogeny from a weighted parsimony analysis (based on PROTPARS weighting) for 19 isocitrate dehydrogenases from animals, plants, fungi, and a proteobacterium. The numbers adjacent to each internal branch indicate the percentages of 1,000 bootstrap replicates that supported the corresponding branch. The two branches marked with asterisks were not supported in any of the equally parsimonious solutions. Abbreviations: cyt, the cytosolic enzyme; mit, the mitochondrial enzyme; Vole m, *M. mexicanus*; Vole o, *M. ochrogaster*; Nematode, *C. elegans*; Yeast ORF, Yeast open reading frame YNL009W; Proteobacterium, *S. yanoikuyae.*

isocitrate dehydrogenases contained a characteristic motif identified by the PROSITE program as the isocitrate and isopropylmalate dehydrogenases signature sequence (fig. 1). This signature sequence is a glycine-rich stretch of amino acid residues common for isocitrate dehydrogenase and 3-isopropylmalate dehydrogenase (Hurley et al. 1989). Second, all mammalian cytosolic ICDHs contain a C-terminal tripeptide, alanine-lysine-leucine, which has been identified by the PROSITE program as a type I peroxisomal targeting signal (PTS1). PTS1 is a tripeptide sequence which is typically found at the C-terminus of peroxisomal proteins (Gould et al. 1989; McNew and Goodman 1996). The tobacco and potato ICDHs also contain a PTS1 tripeptide (alanine-lysine-alanine) at the extreme C-terminus, as well as another PTS1 tripeptide (serine-lysine-leucine) in positions 386–388 (potato) and 385–387 (tobacco). Deduced amino acid sequences of ICDH from soybean (Udvardi, McDermott, and Kahn 1993) and eucalyptus (Boiffin et al., unpublished data) include multiple internal PTS1 sequences. The PSORT program identified the alanine-asparagine-leucine tripeptide as the most probable actual targeting sequence (positions 408–410 in eucalyptus, 409–411 in soybean) for these species. In contrast, the polypeptide from alfalfa (Shorrosh and Dixon 1992) contains only one internal PTS1 tripeptide (serine-lysine-leucine), located at amino acid positions 385–387. The yeast cytosolic enzyme (Loftus et al. 1994) contains the serine-lysine-leucine sequence at amino acid positions 401–403; however, it is not clear whether it is a functional targeting sequence. Our analysis indicates that the nematode polypeptide (Wilson et al. 1994) does not contain a peroxisomal targeting sequence. The physiological importance of the putative peroxisomal targeting sequences found in these isocitrate dehydrogenases is unclear; however, NADP-dependent isocitrate dehydrogenase activity has been reported in rat liver peroxisomes (Leighton et al. 1968).

No subcellular localization has been proposed for the nematode enzyme (Wilson et al. 1994) or for the plant enzymes from alfalfa, soybean, or eucalyptus (Shorrosh and Dixon 1992; Udvardi, McDermott, and Kahn 1993; Boiffin et al., unpublished data). We analyzed amino acid sequences from these taxa for the presence of the mitochondrial and the chloroplastic signal sequences using the PSORT II program and did not identify either signal.

Predicted Isoelectric Points of ICDHs

In order to globally characterize the amino acid composition of the proteins listed in table 2, we calculated their isoelectric point (*p*I) values based on their primary structures using the computer program MacVector (see *Materials and Methods*). The estimation of the *p*I of a protein from its amino acid sequence is not precise. For example, Fatania, Al-Nassar, and Sidhan (1993) used isoelectric focusing to determine the experimental *p*I of the rat cytosolic NADP-dependent isocitrate dehydrogenase, which they found to be 5.7. The *p*I predicted from the amino acid sequence by MacVector is 6.6. This difference is explainable, because MacVector calculates *p*I assuming that the degree of the dissociation of an amino acid in a protein is the same as that of a free amino acid. Despite these potential errors in measurement, all cytosolic ICDHs had an average predicted *p*I value of 6.4 (human, 6.4; mouse, 6.5; *M. mexicanus*, 6.9; *M. ochrogaster*, 6.9; rat, 6.6; nematode, 6.0; alfalfa, 6.2; soybean, 5.8; eucalyptus, 6.5; tobacco, 6.2; potato, 6.9; yeast, 5.8), whereas mitochondrial forms known to date had a basic predicted average *p*-Ivalue of 9.1 (human, 9.1; mouse, 9.2; pig, 8.9; cow, 9.1; yeast, 9.1). The predicted difference in *p*I values for the cytosolic and mitochondrial enzymes may be a result of the different chemical environments of the cytoplasm and the mitochondrion, and may serve as a diagnostic character for distinguishing between amino acid sequences of the cytosolic and mitochondrial ICDH.

Phylogenetic Analysis of the ICDH Family

One of the three shortest trees found by the weighted parsimony analysis of the inferred amino acid sequences of the ICDH genes is shown in figure 2, along with bootstrap support percentages for each of the internal branches. The three most-parsimonious trees differed only as far as minor rearrangements among some of the mammalian mitochondrial enzymes. In all three trees, the mammalian cytosolic enzymes grouped strongly together (100% of the bootstrap replicates), and together formed a group with the nematode enzyme (93% of bootstrap replicates). Likewise, there was strong support for a grouping of mammalian mitochondrial enzymes (100% of bootstrap replicates), a grouping of plant enzymes (100% of bootstrap replicates), and a grouping of all three forms of the enzyme from yeast (100% of bootstrap replicates). All of the most-parsimonious trees also grouped the animal mitochondrial enzymes with the animal cytosolic enzymes, but bootstrap support for this grouping was weak (42%).

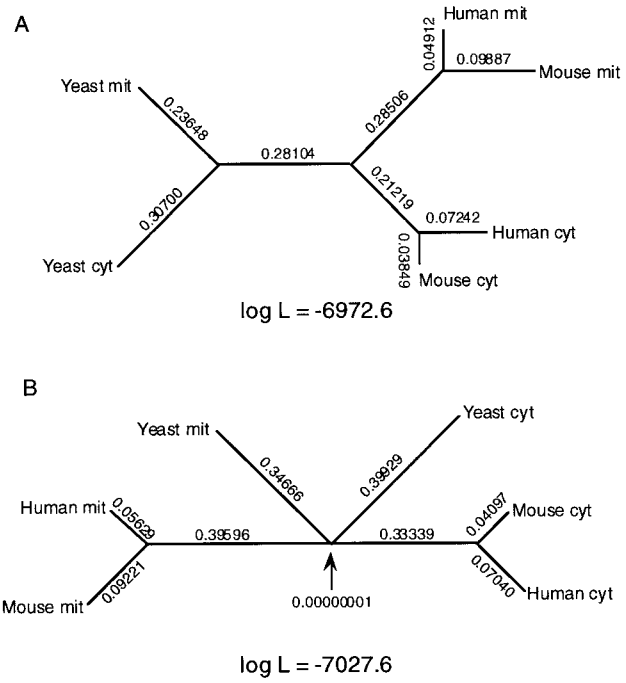

FIG. 3.—Maximum-likelihood analyses based on the nucleotide sequences of three cytosolic and three mitochondrial ICDH genes. *A*, The optimal tree groups the fungal mitochondrial and cytosolic ICDH genes separate from the mammalian mitochondrial and cytosolic ICDH genes, with a long branch separating the two groups. *B*, When the analysis is constrained to force the mitochondrial and cytosolic genes to group together, there is virtually no support for an internal branch separating these two groups, and the resultant tree is 55 log-likelihood units worse. Thus, the hypothesis that the mitochondrial and cytosolic ICDH genes are orthologous can be rejected ($P \ll 0.001$). Abbreviations: cyt, cytosolic; mit, mitochondrial.

A priori, we had assumed that the mitochondrial and cytosolic forms of ICDH were derived from an ancestral gene through a single gene duplication event early in the eukaryotic radiation. If this were true, then we would expect all mitochondrial ICDH sequences to cluster together and all cytosolic ICDH sequences also to cluster together, forming two independent clades. Instead, the parsimony analysis (fig. 2) shows that the mitochondrial forms of ICDH in fungi and mammals are more closely related to the cytosolic forms of the enzyme in these same lineages, suggesting that the genes arose through separate gene duplication events and are not orthologous in animals and fungi. To further investigate the finding that mitochondrial and cytosolic ICDH genes do not appear to be orthologous in animals and fungi, we examined the available nucleotide sequences for these enzymes from three taxonomic groups (humans, mice, and yeast). First, we conducted a likelihood ratio test to compare the best maximum-likelihood tree for these six sequences (fig. 3*A*) to the best tree under the constraint that the genes for the mitochondrial ICDH are orthologous in humans, mice, and fungi and are paralogous to the genes of the cytosolic ICDH (the null hypothesis; fig. 3*B*). The best tree (fig. 3*A*) groups the two yeast genes separate from the four mammalian genes, whereas the null hypothesis assumes that the genes for the three mitochondrial enzymes group sepa-

rately from the genes for the three cytosolic enzymes (fig. 3*B*). There was considerably more support for the best tree (fig. 3*A*; log-likelihood score = −6972.6) than for the null hypothesis (fig. 3*B*; log-likelihood score = −7027.6). As the test statistic δ (two times the difference in log-likelihoods of the two hypotheses; δ = 110 in this test) is distributed approximately as a χ^2 (df = 1) distribution (Felsenstein 1988), the null hypothesis is rejected at $P \ll 0.001$ (the critical value of χ^2 [α = 0.001, df = 1] is 10.828).

The nearly zero length branch that separates the mitochondrial and cytosolic forms of the enzymes in the constrained tree (fig. 3*B*) further indicates that there is essentially no support for the null hypothesis of orthology among mitochondrial forms of ICDH in animals and fungi. We also searched for indications of possible common ancestry for the mitochondrial enzymes by comparing the numbers of apparent homoplastic substitutions (as reconstructed by parsimony) on the best tree (topology as in fig. 3*A*). (Note that we are not trying to estimate convergence; we are looking for evidence that would provide support for the alternative hypothesis.) If, for example, the three mitochondrial ICDH loci actually were orthologous (i.e., the duplication happened before speciation of animals and fungi), but the genes grouped apart on the tree because of partial gene conversion events, we would predict that there would be significantly more homoplastic substitutions between each of the yeast genes and their putative mammalian orthologs than between each of the yeast genes and their alternative mammalian paralogs (e.g., the yeast mitochondrial gene would share more apparently homoplastic substitutions with the mammalian mitochondrial genes than with the mammalian cytosolic genes). We found 60 homoplasies between putative orthologs (mitochondrial vs. mitochondrial, cytosolic vs. cytosolic) and 53 between putative paralogs (mitochondrial vs. cytosolic); these totals are not significantly different based on a binomial test (P = 0.5725). If gene conversion events were complete rather than partial, then the sequence at the site of the gene conversion event would be orthologous to the replacing sequence. In addition, if gene conversion events were common between the two loci in the same organism, we would also expect some indication of homoplastic substitutions between the genes for the human mitochondrial and cytosolic forms of the enzyme or between the genes for the mouse mitochondrial and cytosolic forms of the enzyme. Instead, there are fewer such substitutions (a total of 3 homoplastic events) than between the mouse mitochondrial and human cytosolic forms or between the human mitochondrial and mouse cytosolic forms of the enzyme (a total of 12 homoplastic events). Thus, there is no indication of whole or partial orthology between the fungal and animal mitochondrial forms of ICDH or between the fungal and animal cytosolic ICDHs, suggesting that both forms of the enzyme arose through independent duplication of an ancestral ICDH gene.

Conclusions

In this report, we summarized sequence data available for 18 eukaryotic ICDHs. Four of these sequences (cytosolic ICDH from humans, mice, and two species of voles) were obtained in our laboratory and are described here for the first time. Comparison of structural features and phylogenetic trees of these 18 known ICDH sequences allowed us to conclude that the previously unclassified ICDHs from nematode, alfalfa, soybean, and eucalyptus are most likely cytosolic enzymes (see *Protein Sequence Motifs, Predicted Isolelectric Points of ICDHs,* and *Phylogenetic Analysis of the ICDH Family*).

The most surprising aspect of the phylogenetic analysis was the finding that the yeast mitochondrial and cytosolic enzymes do not appear to be orthologous to their functional counterparts in animals. We had expected that the genes for all the mitochondrial and cytosolic forms of ICDH had descended from a common ancestral gene that duplicated prior to the divergence of the crown eukaryotes and gave rise to two lineages of ICDH genes. However, neither the mitochondrial nor the cytosolic forms of ICDH from mammals and fungi group together. Instead, the yeast mitochondrial form of ICDH appears to be more closely related to the yeast cytosolic form of the enzyme, and the same relationship holds true for the corresponding pairs of ICDH genes in mammals (figs. 2 and 3). Although it is possible that gene conversion events within fungi and within mammals have obliterated the true history of the ICDH gene duplications, the lack of any indication of putative homoplastic substitutions between the mitochondrial forms of the enzyme in fungi and animals provides no support for this hypothesis.

The amino acid sequence of an isocitrate dehydrogenase from an α-proteobacterium (*S. yanoikuyae*) published by Wang and Lau (1996) has an unusually high similarity to the eukaryotic ICDHs (table 2). It also differs remarkably from the *E. coli* enzyme, which is very divergent from the mammalian ICDHs (approximately 14% amino acid identity on average). The fact that *S. yanoikuyae* belongs to the α-subdivision of proteobacteria, the group that gave rise to the mitochondrion through endosymbiosis (Yang et al. 1985; Gray and Spencer 1996), allows us to propose the following scenario of ICDH evolution: First, an α-proteobacterium carrying a precursor ICDH gene was engulfed by a nucleus-containing host cell. Second, this gene was transferred into the nuclear genome, forming an ancestral ICDH locus. Third, two independent duplication events took place in animals and fungi (after separation of these lineages) at the ancestral ICDH locus, giving rise to what we know today as the mitochondrial and cytosolic forms of ICDH. One potential problem with the proposed hypothesis is that the mitochondrial genome from a freshwater protozoon *Reclinomonas americana,* containing the largest number of genes (97) of any known mitochondrial DNA (Lang et al. 1997; Gray et al. 1998), does not contain any ICDH-related sequences. However, we can further speculate that the transfer of a precursor ICDH gene from an early mitochondrial genome to the nucleus of a host cell occurred rapidly after the endosymbiotic event or that this particular endosymbiont did not become a mitochondrion. Martin et al. (1998) pro-

pose that gene transfer between organellar and nuclear genomes is quite favorable, because it increases recombination and reduces genetic load. Additionally, the duplication of the precursor ICDH gene was beneficial, because compartmentalized eukaryotic cells needed more than one enzyme to catalyze similar reactions in the cytoplasm and mitochondrion. For example, NADP(H) cannot penetrate the mitochondrial membrane, and isocitrate is an intramitochondrial source of the reductional equivalent. This requires a specific dehydrogenase, which is the NADP-dependent mitochondrial enzyme. In the cytoplasm, a similar function is performed by the cytosolic enzyme. This function appears to be essential in the eukaryotic cell, as can be seen from the high degree of conservation among cytosolic ICDHs characterized from various organisms.

Acknowledgments

This work is part of the Chernobyl Research Program directed by Dr. Ronald K. Chesser of the Savannah River Ecology Laboratory. Dr. David L. Swofford kindly provided the software program PAUP*. Special thanks to Dr. Susan San Francisco and Ruwanthi Wettasinghe for technical assistance. Suggestions made by the reviewing editor and two anonymous reviewers greatly helped us to improve the quality of the manuscript. We also thank the Field Methods class, Texas Tech University, for assistance in specimen collection; and the Natural Sciences Research Laboratory, the Museum of the Texas Tech University, for loan of tissues used in this study. Financial support was provided by contract DE-FC09–96SR18546 between the U.S. Department of Energy and the University of Georgia and by Texas Tech University.

LITERATURE CITED

BAIROCH, A., P. BUCHER, and K. HOFMANN. 1997. The PROSITE database, its status in 1997. Nucleic Acids Res. **25**: 217–222.

BELFIORE, F., and S. IANNELLO. 1995. Fatty acid synthesis from glutamate in the adipose tissue of normal subjects and obese patients: an enzyme study. Biochem. Mol. Med. **54**:19–25.

CHEN, R. D., E. BISMUTH, M. L. CHAMPIGNY, and P. GADAL. 1989. Chromatographic and immunological evidence that chloroplastic and cytosolic pea leaf NADP-dependent isocitrate dehydrogenase are distinct isozymes. Planta **178**: 157–163.

FATANIA, H., K. E. AL-NASSAR, and V. SIDHAN. 1993. Purification and partial characterization of NADP-linked isocitrate dehydrogenase from rat liver cytosol. FEBS Lett. **320**: 57–60.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. **22**:521–565.

———. 1993. PHYLIP (phylogeny inference package). Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

FIEUW, S., B. MULLER-ROBER, S. GALVEZ, and L. WILLMITZER. 1995. Cloning and expression analysis of the cytosolic NADP(+)-dependent isocitrate dehydrogenase from potato: implications for nitrogen metabolism. Plant Physiol. **107**: 905–913.

GALVEZ, S., M. HODGES, P. DECOTTIGNIES, E. BISMUTH, M. LANCIEN, R. S. SANGWAN, F. DUBOIS, P. LEMARECHAL, C. CRETIN, and P. GADAL. 1996. Identification of a tobacco cDNA encoding a cytosolic NADP-isocitrate dehydrogenase. Plant Mol. Biol. **30**:307–320.

GOULD, J. S., G.-A. KELLER, N. HOSKEN, J. WILKINSON, and S. SUBRAMANI. 1989. A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. **108**:1657–1664.

GRAY, M. W., B. F. LANG, R. CEDERGREN et al. (15 co-authors). 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. **26**:865–868.

GRAY, M. W., and D. F. SPENCER. 1996. Organellar evolution. Pp. 109–126 *in* D. M. ROBERTS, P. SHARP, G. ALDERSON, and M. COLLINS, eds. Evolution of microbial life. Cambridge University Press, Cambridge, England.

HASELBECK, R. J., R. F. COLMAN, and L. MCALISTER-HENN. 1992. Isolation and sequence of cDNA encoding porcine mitochondrial NADP-specific isocitrate dehydrogenase. Biochemistry **31**:6219–6223.

HASELBECK, R. J., and L. MCALISTER-HENN. 1991. Isolation, nucleotide sequence, and disruption of the Saccharomyces cerevisiae gene encoding mitochondrial NADP(H)-specific isocitrate dehydrogenase. J. Biol. Chem. **266**:2339–2345.

HASELBECK, R. J., and L. MCALISTER-HENN. 1993. Function and expression of yeast mitochondrial NAD- and NADP-specific isocitrate dehydrogenases. J. Biol. Chem. **268**: 12116–12122.

HILLIS, D. M., B. K. MABLE, and C. MORITZ. 1996. Applications of molecular systematics: the state of the field and a look to the future. Pp. 515–543 *in* D. M. HILLIS, B. K. MABLE, and C. MORITZ, eds. Molecular systematics. 2nd edition. Sinauer, Sunderland, Mass.

HUELSENBECK, J. P., D. M. HILLIS, and R. NIELSEN. 1996. A likelihood-ratio test of monophyly. Syst. Biol. **45**:546–558.

HUH, T. L., J. H. RYU, J. W. HUH, H. C. SUNG, I. U. OH, B. J. SONG, and R. L. VEECH. 1993. Cloning of a cDNA encoding bovine mitochondrial NADP(+)-specific isocitrate dehydrogenase and structural comparison with its isoenzymes from different species. Biochem. J. **292**:705–710.

HURLEY, J. H., A. M. DEAN, D. E. KOSHLAND JR., and R. M. STROUD. 1991. Catalytic mechanism of $NADP^+$-dependent isocitrate dehydrogenase: implication from the structures of magnesium-isocitrate and $NADP^+$ complexes. Biochemistry **30**:8671–8678.

HURLEY, J. H., A. M. DEAN, J. L. SOHL, D. E. KOSHLAND JR., and R. M. STROUD. 1990. Regulation of an enzyme by phosphorylation at the active site. Science **249**:1012–1016.

HURLEY, J. H., P. E. THORSNESS, V. RAMALINGAM, N. H. HELMERS, D. E. KOSHLAND JR., and R. M. STROUD. 1989. Structure of a bacterial enzyme regulated by phosphorylation, isocitratedehydrogenase. Proc. Natl. Acad. Sci. USA **86**:8635–8639.

JAENICKE, T., K. W. DIEDERICH, W. HAAS, J. SCHLEICH, P. LICHTER, M. PFORDT, A. BACH, and H. P. VOSBERG. 1990. The complete sequence of the human beta-myosin heavy chain gene and a comparative analysis of its product. Genomics **8**:194–206.

JENNINGS, G. T., S. SECHI, P. M. STEVENSON, R. C. TUCKEY, D. PARMELEE, and L. MCALISTER-HENN. 1994. Cytosolic $NADP^+$-dependent isocitrate dehydrogenase: isolation of rat cDNA and study of tissue-specific and developmental expression of mRNA. J. Biol. Chem. **269**:23128–23134.

JOHNSTON, M., L. HILLIER, L. RILES et al. (56 co-authors) 1997. The nucleotide sequence of Saccharomyces cerevisiae chromosome XII. Nature **387**:87–90.

KEYS, D. A., and L. MCALISTER-HENN. 1990. Subunit structure, expression, and function of NAD(H)-specific isocitrate dehydrogenase in Saccharomyces cerevisiae. J. Bacteriol. **172**:4280–4287.

KRAFT, R., M. BRAVO-ZEHNDER, D. A. TAYLOR, and L. A. LEINWAND. 1989. Complete nucleotide sequence of full length cDNA for rat beta cardiac myosin heavy chain. Nucleic Acids Res. **17**:7529–7530.

LANG, B. F., G. BURGER, C. J. O'KELLY, R. CEDERGREN, G. B. GOLDING, C. LEMIEUX, D. SANKOFF, M. TURMEL, and M. W. GRAY. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature **387**: 493–497.

LEIGHTON, F., B. POOLE, H. BEAUFAY, P. BAUDHUIN, J. W. COFFEY, S. FOWLER, and C. DE DUVE. 1968. The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with triton WR-1339. Improved isolation procedures, automated analysis, biochemical and morphological properties of fractions. J. Cell Biol. **37**:482–513.

LI, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, Mass.

LOFTUS, T. M., L. V. HALL, S. L. ANDERSON, and L. MCALISTER-HENN. 1994. Isolation, characterization and disruption of the yeast gene encoding cytosolic NADP-specific isocitrate dehydrogenase. Biochemistry **33**:9661–9667.

MCALISTER-HENN, L., and W. C. SMALL. 1997. Molecular genetics of yeast TCA cycle isozymes. Prog. Nucleic Acid Res. Mol. Biol. **57**:317–339.

MCNEW, J. A., and J. M. GOODMAN. 1996. The targeting and assembly of peroxisomal proteins: some old rules do not apply. Trends Biochem. Sci. **21**:54–58.

MARTIN, W., B. STOEBE, V. GOREMYKIN, S. HANSMANN, M. HASEGAVA, and K. V. KOVALIK. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature **393**:162–165.

NAKAI, K. 1997. PSORT II. World Wide Web: http://psort.nibb.ac.jp/psort

RAMACHANDRAN, N., and R. F. COLMAN. 1980. Chemical characterization of distinct subunits of pig heart NAD-specific isocitrate dehydrogenase. J. Biol. Chem. **255**:8859–8864.

SHORROSH, B. S., and R. A. DIXON. 1992. Molecular characterization and expression of an isocitrate dehydrogenase from alfalfa (*Medicago sativa* L). Plant Mol. Biol. **20**:801–807.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (and other methods). Version 4.0. Sinauer, Sunderland, Mass.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. HILLIS, B. K. MABLE, and C. MORITZ, eds. Molecular systematics. 2nd edition. Sinauer, Sunderland, Mass.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

UDVARDI, M. K., T. R. MCDERMOTT, and M. L. KAHN. 1993. Isolation and characterization of a cDNA encoding NADP(+)-specific isocitrate dehydrogenase from soybean (*Glycine max*). Plant Mol. Biol. **21**:739–752.

WAHLE, E., and W. KELLER. 1996. The biochemistry of polyadenylation. Trends Biochem. Sci. **21**:247–250.

WANG, Y., and P. C. K. LAU. 1996. Sequence and expression of an isocitrate dehydrogenase-encoding gene from a polycyclic aromatic hydrocarbon oxidizer, *Sphingomonas yanoikuyae* B1. Gene **168**:15–21.

WILSON, R., R. AINSCOUGH, K. ANDERSON et al. (53 co-authors). 1994. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans.* Nature **368**:32–38.

WINKLER, B. S., N. DESANTIS, and F. SOLOMON. 1986. Multiple NADPH-producing pathways control glutathione content in retina. Exp. Eye Res. **43**:829–847.

YAN, B., D. YANG, and A. PARKINSON. 1995. Cloning and expression of hydrolase C, a member of the rat carboxylesterase family. Arch. Biochem. Biophys. **317**:222–234.

YANG, L., H. LUO, P. VINAY, and J. WU. 1996. Molecular cloning of the cDNA of mouse mitochondrial NADP-dependent isocitrate dehydrogenase and the expression of the gene during lymphocyte activation. J. Cell. Biochem. **60**:400–410.

YANG, D., Y. OYAIZU, H. OYAIZU, G. J. OLSEN, and C. R. WOESE. 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82**:4443–4447.

ZHAO, W.-N., and L. MCALISTER-HENN. 1996. Expression and gene disruption analysis of the isocitrate dehydrogenase family in yeast. Biochemistry **35**:7873–7878.

CARO-BETH STEWART, reviewing editor

Accepted September 15, 1998