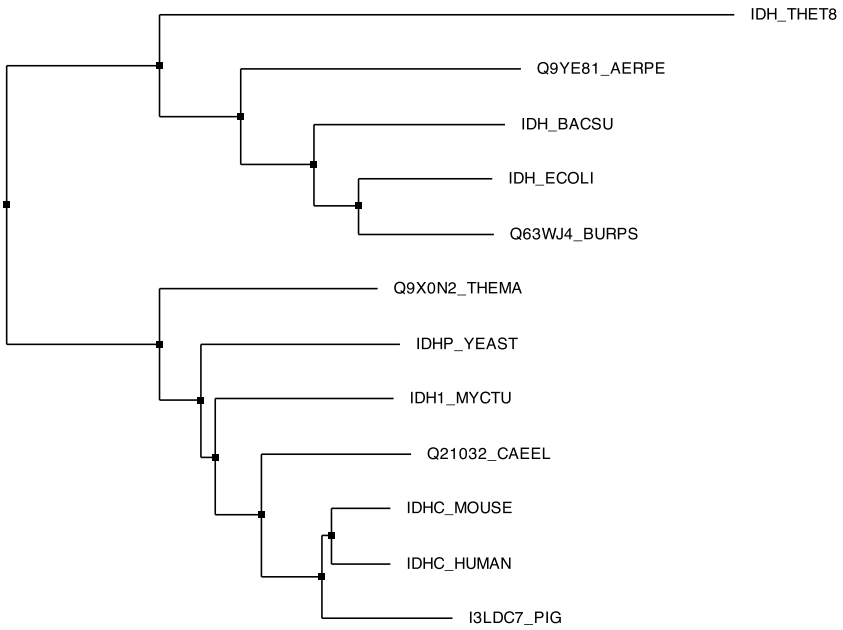
Figure S1: Sequence Analysis of IDH-1

A



B

	T. MAR.	*S. CER.*	*M. TUB.*	*C. ELE.*	*M. MUS.*	*H. SAP.*	*S. SCRO.*	*T. THERM.*	*A. PER.*	*B. SUB.*	*E. COLI*	*B. PSEUD.*
Q9X0N2_THEMA	100	56.2	55.5	51.1	53.1	53.4	52.4	22.1	18.6	23.6	23.4	24.2
IDHP_YEAST	56.2	100	61.9	62.8	65.3	65.5	62.8	21.7	18.0	19.8	17.2	19.7
IDH1_MYCTU	55.5	61.9	100	64.9	65.0	65.5	64.3	21.8	20.7	20	21.9	19.9
Q21032_CAEEL	51.1	62.8	64.9	100	76.5	75.8	75.9	21.3	18.1	20	19.8	20
IDHC_MOUSE	53.1	65.3	65.0	76.5	100	95.7	95.2	22.0	19.8	20	18.1	18.9
IDHC_HUMAN	53.4	65.5	65.5	75.8	95.7	100	95.7	21.7	19.3	19.4	17.6	18.9
I3LDC7_PIG	52.4	62.8	64.3	75.8	95.2	95.7	100	21.6	18.6	19.4	17.8	18.7
IDH_THET8	22.1	21.7	21.8	21.3	22.0	21.7	21.6	100	37.1	39.2	39.7	36.6
Q9YE81_AERPE	18.6	18.0	20.7	18.1	19.8	19.3	18.6	37.1	100	48.9	48.2	44.6
IDH_BACSU	23.6	19.8	20	20	20	19.4	19.4	39.2	48.9	100	68.8	67.4
IDH_ECOLI	23.4	17.2	21.9	19.8	18.1	17.6	17.8	39.7	48.2	68.8	100	74.8
Q63WJ4_BURPS	24.2	19.7	19.9	20	18.9	18.9	18.7	36.6	44.6	67.4	74.8	100

C


Q9X0N2_THEMA/1-399
IDHP_YEAST/1-428
IDH1_MYCTU/1-409
Q21032_CAEEL/1-412
IDHC_MOUSE/1-414
IDHC_HUMAN/1-414
I3LDC7_PIG/1-463
IDH_THET8/1-496
Q9YE81_AERPE/1-432
IDH_BACSU/1-423
IDH_ECOLI/1-416
Q63WJ4_BURPS/1-419

C continued

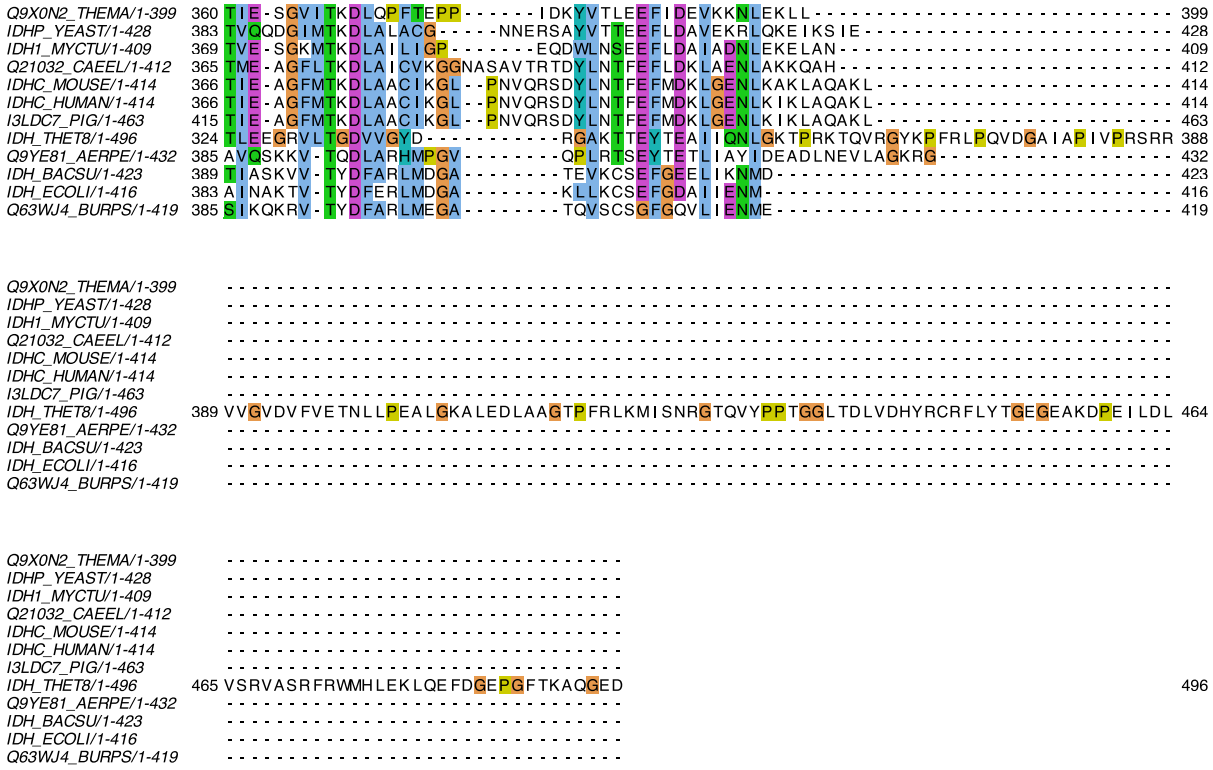


Figure S1. Protein sequence analysis of IDH-1. A. Phylogenetic tree of *C. elegans* IDH-1 (Q21032_CAEEL) is shown with $NADP^+$-dependent IDH enzymes from three IDH subfamilies. Subfamily I consists of dimeric bacterial enzymes (Q9YE81_AERPE, *Aeropyrum pernix*; IDH_BACSU, *Bacillus subtilis*; IDH_ECOLI, *Escherichia coli*; Q63WJ4_BURPS, *Burkholderia pseudomallei*). Subfamily II consists of dimeric eukaryotic enzymes and some dimeric bacterial enzymes (Q9X0N2_THEMA, *Thermotoga maritima*; IDHP_YEAST, *Saccharomyces cerevisiae*; IDH1_MYCTU, *Mycobacterium tuberculosis*; Q21032_CAEEL, *C. elegans*; IDHC_MOUSE, *Mus musculus*; IDHC_HUMAN, *Homo sapiens*; I3LDC7_PIG, *Sus scrofa*). Subfamily IV consists of longer dimeric proteins with larger clasp regions (IDH_THET8, *Thermus thermophilus*). Subfamily III enzymes bind to NAD^+ and are not included. The names in the diagram are the UniProt IDs. B. Percent sequence identity matrix for the IDH sequences shown in A and C. Yellow highlights Subfamily I, and green highlights Subfamily II. The UniProt IDs are shown on the left, and the genus and species abbreviations are shown on the top. C. Sequence alignment of the IDH sequences shown in A and B. The proteins are named using their UniProt IDs. The amino acids are colored using a ClustalW coloring scheme. The sequence alignment and percent identity matrix were determined using Clustal Omega at EMBL-EBI. The phylogenetic tree was determined from the alignment in JalView using neighbor joining with BLOSUM62.