Homo sapiens aldehyde dehydrogenase 2 family member (ALDH2), transcript variant 1, mRNA; nuclear gene for mitochondrial product

NCBI Reference Sequence: NM_000690.4

FASTA Graphics

Go to:

LOCUS NM_000690 9561 bp mRNA linear PRI 27-APR-2025

DEFINITION Homo sapiens aldehyde dehydrogenase 2 family member (ALDH2), transcript variant 1, mRNA; nuclear gene for mitochondrial product.

ACCESSION NM_000690

VERSION NM_000690.4

KEYWORDS RefSeq; MANE Select.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 9561)

AUTHORS Yokoyama,A., Yokoyama,T., Yumoto,Y., Takimura,T., Toyama,T., Yoneda,J., Nishimura,K., Minobe,R., Matsuzaki,T., Kimura,M. and Matsushita,S.

TITLE Associations of ADH1B and ALDH2 genotypes and alcohol flushing with drinking history, withdrawal symptoms, and ICD-10 criteria in Japanese alcohol-dependent men

JOURNAL Pharmacogenet Genomics 34 (5), 139-148 (2024)

PUBMED 38465575

REMARK GeneRIF: Associations of ADH1B and ALDH2 genotypes and alcohol flushing with drinking history, withdrawal symptoms, and ICD-10 criteria in Japanese alcohol-dependent men.

REFERENCE 2 (bases 1 to 9561)

AUTHORS Liu,Z.Y., Lin,X.H., Guo,H.Y., Shi,X., Zhang,D.Y., Sun,J.L., Zhang,G.C., Xu,R.C., Wang,F., Yu,X.N., Wang,D., Weng,S.Q., Shen,X.Z., Liu,T.T., Dong,L. and Zhu,J.M.

TITLE Multi-Omics profiling identifies aldehyde dehydrogenase 2 as a critical mediator in the crosstalk between Treg-mediated immunosuppression microenvironment and hepatocellular carcinoma

JOURNAL Int J Biol Sci 20 (7), 2763-2778 (2024)

PUBMED 38725845

REMARK GeneRIF: Multi-Omics profiling identifies aldehyde dehydrogenase 2 as a critical mediator in the crosstalk between Treg-mediated immunosuppression microenvironment and hepatocellular carcinoma. Publication Status: Online-Only

REFERENCE 3 (bases 1 to 9561)

AUTHORS Wang,X., Wang,J., Chen,Y., Qian,X., Luo,S., Wang,X., Ma,C. and Ge,W.

TITLE The aldehyde dehydrogenase 2 rs671 variant enhances amyloid beta

pathology
JOURNAL Nat Commun 15 (1), 2594 (2024)
PUBMED 38519490
REMARK GeneRIF: The aldehyde dehydrogenase 2 rs671 variant enhances amyloid beta pathology.
Publication Status: Online-Only
REFERENCE 4 (bases 1 to 9561)
AUTHORS Liu,R., Peng,M., Zhang,J., Qiu,K., Zeng,T. and Chen,L.
TITLE The ALDH2 gene rs671 polymorphism is associated with cardiometabolic risk factors in East Asian population: an updated meta-analysis
JOURNAL Front Endocrinol (Lausanne) 15, 1333595 (2024)
PUBMED 38567307
REMARK GeneRIF: The ALDH2 gene rs671 polymorphism is associated with cardiometabolic risk factors in East Asian population: an updated meta-analysis.
Publication Status: Online-Only
REFERENCE 5 (bases 1 to 9561)
AUTHORS Tokiya,M., Hashimoto,M., Fukuda,K., Kawamoto,K., Akao,C., Tsuji,M., Yakushiji,Y., Koike,H. and Matsumoto,A.
TITLE Asian flush gene variant increases mild cognitive impairment risk: a cross-sectional study of the Yoshinogari Brain MRI Checkup Cohort
JOURNAL Environ Health Prev Med 29, 55 (2024)
PUBMED 39401906
REMARK GeneRIF: Asian flush gene variant increases mild cognitive impairment risk: a cross-sectional study of the Yoshinogari Brain MRI Checkup Cohort.
REFERENCE 6 (bases 1 to 9561)
AUTHORS Yoshida,A.
TITLE Molecular genetics of human aldehyde dehydrogenase
JOURNAL Pharmacogenetics 2 (4), 139-147 (1992)
PUBMED 1306115
REMARK Review article
REFERENCE 7 (bases 1 to 9561)
AUTHORS Crabb,D.W., Edenberg,H.J., Bosron,W.F. and Li,T.K.
TITLE Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant
JOURNAL J Clin Invest 83 (1), 314-316 (1989)
PUBMED 2562960
REFERENCE 8 (bases 1 to 9561)
AUTHORS Hsu,L.C., Bendel,R.E. and Yoshida,A.
TITLE Genomic structure of the human mitochondrial aldehyde dehydrogenase gene
JOURNAL Genomics 2 (1), 57-65 (1988)
PUBMED 2838413

REFERENCE 9 (bases 1 to 9561)
AUTHORS Hsu,L.C., Tani,K., Fujiyoshi,T., Kurachi,K. and Yoshida,A.
TITLE Cloning of cDNAs for human aldehyde dehydrogenases 1 and 2
JOURNAL Proc Natl Acad Sci U S A 82 (11), 3771-3775 (1985)
PUBMED 2987944
REFERENCE 10 (bases 1 to 9561)
AUTHORS Greenfield,N.J. and Pietruszko,R.
TITLE Two aldehyde dehydrogenases from human liver. Isolation via affinity chromatography and characterization of the isozymes
JOURNAL Biochim Biophys Acta 483 (1), 35-45 (1977)
PUBMED 18196
COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from DC353712.1, BC002967.1, CB851683.1 and AC003029.5.
This sequence is a reference standard in the RefSeqGene project.

On Nov 22, 2018 this sequence version replaced NM_000690.3.

Summary: This protein belongs to the aldehyde dehydrogenase family of proteins. Aldehyde dehydrogenase is the second enzyme of the major oxidative pathway of alcohol metabolism. Two major liver isoforms of aldehyde dehydrogenase, cytosolic and mitochondrial, can be distinguished by their electrophoretic mobilities, kinetic properties, and subcellular localizations. Most Caucasians have two major isozymes, while approximately 50% of East Asians have the cytosolic isozyme but not the mitochondrial isozyme. A remarkably higher frequency of acute alcohol intoxication among East Asians than among Caucasians could be related to the absence of a catalytically active form of the mitochondrial isozyme. The increased exposure to acetaldehyde in individuals with the catalytically inactive form may also confer greater susceptibility to many types of cancer. This gene encodes a mitochondrial isoform, which has a low Km for acetaldehydes, and is localized in mitochondrial matrix. Alternative splicing results in multiple transcript variants encoding distinct isoforms.[provided by RefSeq, Nov 2016].

Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (1).

Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.

##Evidence-Data-START##

Transcript exon combination :: SRR1803611.169202.1,
SRR1803612.222503.1 [ECO:0000332]
RNAseq introns :: single sample supports all introns
SAMEA1965299, SAMEA1966682
[ECO:0000348]
##Evidence-Data-END##

##RefSeq-Attributes-START##
gene product(s) localized to mito. :: reported by MitoCarta
MANE Ensembl match :: ENST00000261733.7/
ENSP00000261733.2
RefSeq Select criteria :: based on conservation,
expression, longest protein
##RefSeq-Attributes-END##
COMPLETENESS: full length.

PRIMARY	REFSEQ_SPAN	PRIMARY_IDENTIFIER	PRIMARY_SPAN	COMP
	1-6	DC353712.1	47-52	
	7-1938	BC002967.1	3-1934	
	1939-2014	CB851683.1	17-92	c
	2015-9561	AC003029.5	18768-26314	

FEATURES Location/Qualifiers
source 1..9561
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.12"
gene 1..9561
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="aldehyde dehydrogenase 2 family member"
/db_xref="GeneID:217"
/db_xref="HGNC:HGNC:404"
/db_xref="MIM:100650"
exon 1..164
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/inference="alignment:Splign:2.1.0"
misc_feature 39..41
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="upstream in-frame stop codon"
CDS 51..1604
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"

/EC_number="1.2.1.3"
/note="isoform 1 precursor is encoded by transcript variant 1; acetaldehyde dehydrogenase 2; nucleus-encoded mitochondrial aldehyde dehydrogenase 2; liver mitochondrial ALDH; ALDH class 2; aldehyde dehydrogenase 2 family (mitochondrial); epididymis secretory sperm binding protein"
/codon_start=1
/product="aldehyde dehydrogenase, mitochondrial isoform 1 precursor"
/protein_id="NP_000681.2"
/db_xref="CCDS:CCDS9155.1"
/db_xref="GeneID:217"
/db_xref="HGNC:HGNC:404"
/db_xref="MIM:100650"
/translation="MLRAAARFGPRLGRRLLSAAATQAVPAPNQQPEVFCNQIFINNE
WHDAVSRKTFPTVNPSTGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRG
RLLNRLADLIERDRTYLAALETLDNGKPYVISYLVDLDMVLKCLRYYAGWADKYHGKT
IPIDGDFFSYTRHEPVGVCGQIIPWNFPLLMQAWKLGPALATGNVVVMKVAEQTPLTA
LYVANLIKEAGFPPGVVNIVPGFGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGS
SNLKRVTLELGGKSPNIIMSDADMDWAVEQAHFALFFNQGQCCCAGSRTFVQEDIYDE
FVERSVARAKSRVVGNPFDSKTEQGPQVDETQFKKILGYINTGKQEGAKLLCGGGIAA
DRGYFIQPTVFGDVQDGMTIAKEEIFGPVMQILKFKTIEEVVGRANNSTYGLAAAVFT
KDLDKANYLSQALQAGTVWVNCYDVFGAQSPFGGYKMSGSGRELGEYGLQAYTEVKTV
TVKVPQKNS"
transit_peptide 51..101
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
misc_feature 75..122
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="propagated from UniProtKB/Swiss-Prot (P05091.2); Region: SIFI-degron.
/evidence=ECO:0000269|PubMed:38297121"
mat_peptide 102..1601
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/product="aldehyde dehydrogenase, mitochondrial isoform 1"
misc_feature 204..206
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 267..269

/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 282..284
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 525..527
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 606..608
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="Transition state stabilizer.
/evidence=ECO:0000250|UniProtKB:P20000; propagated from
UniProtKB/Swiss-Prot (P05091.2); other site"
misc_feature 1152..1154
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 1197..1199
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 1326..1328
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.
/evidence=ECO:0000250|UniProtKB:P47738; propagated from
UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature 1332..1334
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="N6-acetyllysine.

```
                /evidence=ECO:0000250|UniProtKB:P47738; propagated from
                UniProtKB/Swiss-Prot (P05091.2); acetylation site"
misc_feature    1401..1403
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /note="N6-acetyllysine.
                /evidence=ECO:0000250|UniProtKB:P47738; propagated from
                UniProtKB/Swiss-Prot (P05091.2); acetylation site"
exon            165..269
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            270..410
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            411..490
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            491..602
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            603..731
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            732..845
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            846..948
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            949..1133
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            1134..1298
                /gene="ALDH2"
                /gene_synonym="ALDH-E2; ALDHI; ALDM"
                /inference="alignment:Splign:2.1.0"
exon            1299..1456
```

/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/inference="alignment:Splign:2.1.0"

exon 1457..1571
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/inference="alignment:Splign:2.1.0"

exon 1572..9561
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/inference="alignment:Splign:2.1.0"

regulatory 1988..1993
/regulatory_class="polyA_signal_sequence"
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="hexamer: ATTAAA"

polyA_site 2009
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="major polyA site"

regulatory 9541..9546
/regulatory_class="polyA_signal_sequence"
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"
/note="hexamer: ATTAAA"

polyA_site 9561
/gene="ALDH2"
/gene_synonym="ALDH-E2; ALDHI; ALDM"

ORIGIN

```
        1 gagaccctag ctctgctctc ggtccgctcg ctgtccgcta gcccgctgcg atgttgcgcg
       61 ctgccgcccg cttcgggccc cgcctgggcc gccgcctctt gtcagccgcc gccacccagg
      121 ccgtgcctgc ccccaaccag cagcccgagg tcttctgcaa ccagattttc ataaacaatg
      181 aatggcacga tgccgtcagc aggaaaacat tccccaccgt caatccgtcc actggagagg
      241 tcatctgtca ggtagctgaa ggggacaagg aagatgtgga caaggcagtg aaggccgccc
      301 gggccgcctt ccagctgggc tcaccttggc gccgcatgga cgcatcacac aggggccggc
      361 tgctgaaccg cctggccgat ctgatcgagc gggaccggac ctacctggcg gccttggaga
      421 ccctggacaa tggcaagccc tatgtcatct cctacctggt ggatttggac atggtcctca
      481 aatgtctccg gtattatgcc ggctgggctg ataagtacca cgggaaaacc atccccattg
      541 acggagactt cttcagctac acacgccatg aacctgtggg ggtgtgcggg cagatcattc
      601 cgtggaattt cccgctcctg atgcaagcat ggaagctggg cccagccttg gcaactggaa
      661 acgtggttgt gatgaaggta gctgagcaga cacccctcac cgccctctat gtggccaacc
      721 tgatcaagga ggctggcttt ccccctggtg tggtcaacat tgtgcctgga tttggcccca
      781 cggctggggc cgccattgcc tcccatgagg atgtggacaa agtggcattc acaggctcca
      841 ctgagattgg ccgcgtaatc caggttgctg ctgggagcag caacctcaag agagtgacct
      901 tggagctggg ggggaagagc cccaacatca tcatgtcaga tgccgatatg gattgggccg
```

961 tggaacaggc ccacttcgcc ctgttcttca accagggcca gtgctgctgt gccggctccc
1021 ggaccttcgt gcaggaggac atctatgatg agtttgtgga gcggagcgtt gcccgggcca
1081 agtctcgggt ggtcgggaac ccctttgata gcaagaccga gcaggggccg caggtggatg
1141 aaactcagtt taagaagatc ctcggctaca tcaacacggg gaagcaagag ggggcgaagc
1201 tgctgtgtgg tgggggcatt gctgctgacc gtggttactt catccagccc actgtgtttg
1261 gagatgtgca ggatggcatg accatcgcca aggaggagat cttcgggcca gtgatgcaga
1321 tcctgaagtt caagaccata gaggaggttg ttgggagagc caacaattcc acgtacgggc
1381 tggccgcagc tgtcttcaca aaggatttgg acaaggccaa ttacctgtcc caggccctcc
1441 aggcgggcac tgtgtgggtc aactgctatg atgtgtttgg agcccagtca ccctttggtg
1501 gctacaagat gtcggggagt ggccgggagt tgggcgagta cgggctgcag gcatacactg
1561 aagtgaaaac tgtcacagtc aaagtgcctc agaagaactc ataagaatca tgcaagcttc
1621 ctcccctcagc cattgatgga aagttcagca agatcagcaa caaaaccaag aaaaatgatc
1681 cttgcgtgct gaatatctga aaagagaaat ttttcctaca aaatctcttg ggtcaagaaa
1741 gttctagaat ttgaattgat aaacatggtg ggttggctga gggtaagagt atatgaggaa
1801 ccttttaaac gacaacaata ctgctagctt tcaggatgat ttttaaaaaa tagattcaaa
1861 tgtgttatcc tctctctgaa acgcttccta taactcgagt ttatagggga agaaaaagct
1921 attgtttaca attatatcac cattaaggca actgctacac cctgctttgt attctgggct
1981 aagattcatt aaaaactagc tgctcttaac ttactgggta actctttgtc catctcaata
2041 accaaaggaa ggtgggcatg gtggctcacg cctgtaatcc cagcactttg ggaagccgag
2101 gcgggtggat cactcgaggt caggagtttg agaccagcct ggccaacgta gtgaaaccct
2161 gtctctcact aaaaatacac aaattaggca gtcatggtgg ctggcacctg taatcccagc
2221 tacttgggag gctgaggcag gaaaattggt tgaacctggg aggtggaggt tgcagtgagt
2281 ggagatcatg ccactgcact ccagcctggg agtcagagcg agactccatc tcaaaaaaag
2341 taaataaatg aaatagccta aggaaatgct ggcaggcagt aatgatatgg cacactggat
2401 atgatttctg cccctcctct gctgtgggta aacagcttct gtttcatgca tttacttttt
2461 tattttaatt acaccaataa gaatgtgctt gaatgtttca tgcatttaat ttgtttcatg
2521 catttaactt ctggcctgtt tacctgtagc aggctctgag atgtgctcta gagacagaac
2581 ctaaacatga tgcccaagag aggggagaga tactgaatgt ccaatgttct caaatttttt
2641 ttttttttt ttttgagaca gggtcttgct ctgccatcca ggctggagtg cagtggcacg
2701 atcatagctc actgctgcct cgacctcctg ggctcaaatg atcctcctgc ttcagcttcc
2761 tgggctaatg tttgtatttt ttgtggagat gggactttgt tatgttgccc aggctgtttt
2821 caaactcctg gactcaagtg atcctcccgc cttagcctcc taaggtgctg ggattacagg
2881 tgtgagccac tgtgcccaat tcagttctca gtaaaaactt gatcaggcca aatgaaagtg
2941 acatgggcca atttcataat agattgattt ggaaaccaga tgtcaattta cttactttaa
3001 gtgagatctg tatcttcagt aaccagacct cttgagtctc ccacccagaa cactatcttc
3061 ctctagtcaa gatacctggc aaaggccggg cacagtggct cacgcctata attccagcac
3121 tttgggaggc caaggtgggt gggtcacctg aggttgggag ttcgagacca gcctgaccaa
3181 cagggagaaa ccccatctct actgaaaata caaaaaaaaa aaaaaaaact agccgggcat
3241 ggtggctcat gcctgtaatc ccagctactc aggaggctga ggcaggagaa ttgcttgaac
3301 ctgggaggca gaggttgcag tgagccaaga ttgctccatt gtactccagc ctgggcaaca
3361 agagagaaac tccgtctcaa aatataaata aataagaaaa agaagaaatg gaagctgggt
3421 tcaagagggc aatcagcagt ctctaccatc atgtgaaaca ccactctgtg atttgagtct
3481 tgacaatgta ttttcttttt ttttttcttc ttttttttt gagactgagt ctcggtctgt
3541 cgcccaggct ggagtgcagt gacgtgatct tggctcactg caacctctgc ctcctgggtt
3601 ggagcaattc tcctgcctca gcctcctgag tagctgggat tacaggtgtg cgccaccatg

3661 cctggctaat ttttgtattt ttagtagaga cggggttttg ccatgttggc caggctggtg
3721 ttgaaggcct gacctcaggt aatccaccca cctctgcctc ccagagtact gggattacaa
3781 gtgtgagcca ccgtgcccgg ccaactgtag ggaccagccc cacagggtcg gtgggttttt
3841 ctccccatgt gcggagacga gagattgtag aaataaagac acaagacaaa gaggtaaaag
3901 aaaagacagc tgggcttggg ggaccactac caccaagagg cggagaccgg tagaggcccc
3961 gaatgtctgg ctgcactgtt atttattgga tacaaagcaa aaggggcagg gtaaagagtg
4021 tgagtcatct ccaatgacag gtaaggtcac gtgggtcacg tgtccactgg acagggggcc
4081 cttccctgcc tggcagccaa ggcagagata gacagggaga gagagagaca gcttacgcca
4141 ttatttctgc gtatcagaga cttttagtac tttcactaat tttgctactg ttatctaaaa
4201 ggcagagcca ggtgtacagg atggaacatg aaggcagact aggagcgtga ccactgaagc
4261 acagcaccac agggagacgg ttaggcctcc ggataactaa tgattaatga tattcatata
4321 taatcatgtc tgtgatctag atctagtata actcttgttg ttttatatat tttattatac
4381 tggaacagct cgtgccctcg gtctcttgcc tcagcacctg gatggcttgc cgcccacagc
4441 caacaatgtg ttttcaatat gctccccaaa ctacagcaaa ctcctcttaa ctcgggattt
4501 agggtgatgc cctacttatg cctaacttta tcaccagctt aaccctcatc tcattatata
4561 ggatgatatt acttctagga agattttcta tgtatggagg gtgcatcatt taagagtatt
4621 ttaggccagg cactgtggct catgcctgta atcccagcat tttgggaagc tgaggtgggt
4681 gatttgcttc agcccacgag ttcaagatca gcctggccaa tatggtgaaa ccccatctct
4741 acaaaaaaat acaaaaatta gctgagcgtg gtgatgtgtg cctgtggtcc cagttactct
4801 ggaggctgat gtgggaggat cacttgagcc caggaggtca aggctgcagt gagctgtgat
4861 tgcaccactt cactccagcc cagagtgaga ccccatctct gaaaaaaaca caaaaaacaa
4921 gaaaagaata ctgcaaagaa ccttggtagg ggtgaattgg aaaaagtaga gaaatgtaat
4981 atttctttcc tgctgtactc actgtaactg tgagaggatc ggctctttta accaacagat
5041 aagaaaggaa atattagcta tgaagaaatg tctatcaagt tgattttgac aacctagaca
5101 cattaatctc cttgcacttt tagtgacagg aacccattag aaccttaagc agctgtctga
5161 agtaaaaaca aagctctcag gtgatgaacc acctaatctc cctgggtatt cggatgtggg
5221 tcagtcacag gagtattagg tgtaaagccg ctgcttctgg gtagagcctc cacttgtggc
5281 tgcacagagg gcacagagag cttttccttc agcccggaga agcttcctaa gtgtgggcgc
5341 caccttgtgg gacatacatt gaactacgac gctatttctc caagaattaa caggacagcc
5401 cctcccaccg gtgtccttcc aacgatgttg catacaattt ggctgtacat ttggcaggca
5461 gtgtaatgag ggaatcaaag accaggatca catactggaa agaattgaaa acagagactc
5521 ccaacaaaaa cttgtacacg gacgttcaca gcagcactat tcacgacagc tcagagatgg
5581 aatcaaccca aatatcagtc aatagatgaa tggacagaca aaatgtgaca tatccatgca
5641 gtggaatatt atttagccat aaaaaggaag gaagtgctga tgcatgttgc gatgtggatg
5701 aaccttgaaa acatgatgct gagtgaaaga agccagacac aaaaggagaa ggactgtatg
5761 atttcactta taggaaatat ccagaacagg caaatctata tagagagaag gtcgattgct
5821 gggggccctg ggaagggacc aagggggagt gacttctaat ggtggggtgc ccttttgggg
5881 tgatgaaaat gttttggaac cagatacagg tggtggttgt acaacattgt gaatttacta
5941 aatgcaacca aattgctcac ttaaaaacaa caaattcagg ccaggtgcag tggctcacac
6001 ccgtaatccc agcactttgg gaggccgagg tgggcggatc acatgaggtc aagagttcga
6061 gactagcctg gccaacatgg tgaaatccca tctcggccgg gcgcggtggc tcacgcctgt
6121 aatcccagca ctttgggagg ctgaggcggg tggatcacga ggtcaggaga tcgagaccat
6181 cctggctaac acggtgaaac cccgtctcta ctaaaaatac aaaaaattag ccaggcgtgg
6241 tggtgggcgc ctgtagtccc agctgatcgg gaggctaagg caggagaatg gtgtgaaccc
6301 aggaggcgga gcttgcagtg agccgagatc gcgccactgc actctatcct gggtgacaga

6361 gtgagactct gtctccaaaa aaaaaaaaag aaaccccatc tctacccaaa atacaaaaat
6421 tagctgggca tggtggctgg cacctgtaat cccagctact caggaggctg aggcaggaga
6481 attggttgaa cctgggaggc agaggttgca gtgagctaag atcgcaccac tgcactccag
6541 cctgggtgac agaacgagac tgtctcaaaa aaaaaaaaaa aaaaaagtgg ccaggtggtg
6601 ccatggctca tgcctgtaat ctcagcattt tgggatactg aggcgggtgg attacctgag
6661 gccaggagtt cgagatcagc ctggccaaca tgatgaaaac ccatctctgc taaaaataca
6721 aaaaattggc tgggcttggt agcatgtgcc cagctacttg ggaggctgag acaggagaat
6781 cgcttgaacc tgggaggcag aggttgcggt tagccaagat tgtgccactg cactccatgg
6841 gtgacacaga gagactctgt ctcaaaaaaa aaaaaaaaaa agtaaatgct atgttatgtg
6901 aatttcacct caactgaaaa aaacagggtc ccagagctgg tcaggggaac ttggacaaac
6961 cacacaactt ctgagcttcc atttcctcac aatctctgca gtcacttcca actctcattt
7021 ggcttattct gtgaaggata aaacagagct gacatttgtt catttgtatt gattttattt
7081 acatttttca aaaggaaata gatgcagatg gtatagaaag tttcagtgaa atgtaaatct
7141 gatgcccaat cctgtggccc caagaagtat taactgtgca cccactacat gccaggaact
7201 gtacttggag ccttctcaag agcttcttaa caatattttc atttttttata ttttttaatg
7261 acagggtcta actccgtcgt ccaggctgga gtgcagtggt gcgatcatag ctcactgcag
7321 cctctacctc ctaggctcag gtgattcttc tgcctcagcc tcccaaatag ctggtactac
7381 aagcatgtgc caccacagtc agctaattat ttttggagag tcagggtttc actatgtagc
7441 tcaggctggt cttcagctcc tggcttgaag tgatcctcct gccttggcac atagaagggg
7501 ttacaatata catatattag agtgagagca tagatgtgat gcctggcgat atggcttact
7561 tatgagacag cacctgacct acagccatct atgactccta ctgtatcttc catgccagga
7621 gtaagggtga ggatacctgg caataataaa aataataacc tgagagtcta atggcttatt
7681 ttttgtagag atggggtctc actgtgttgt ttaggctggt cttgaacttc tgggctcaag
7741 tgatcctccc acctcagcct cctgagtagc tgggattata ggcgtgcact accatgcata
7801 tatatatata ttttttttttt tttttgagag agagagagag agatggagtc ttgctatgtt
7861 gctcaggctg gtctcgaact cctggcctca agtcatcctg ctggcttggc ctctcaagtg
7921 ctaggattac aggcatgaga cactgggcac agccctaata gcttattttt aaagctatat
7981 caaaatgtca tgcacctgaa catttgaagt aacctctctg atagtacgga tatttatccc
8041 ctcaaaatct catgttaaaa tgtaatcttc agtgttggag gtggggcaag gtgtgaggtg
8101 ttaggatcat gggggtggat ccctcatgaa tggcttggtg tcctccccca tggtaatgag
8161 ttcactcaag agctggttgt ttaaaaggtg acgtggacct aatccctgtt ggactgaaca
8221 aaaggggacg aacacagcaa taaagataaa gaccaaagag tatatttgga agaaggggtc
8281 agggggctcc ttgcttctag tgaaaaaggg ccctgagctt tagtgccctt tgtatttatt
8341 gggtaaagga gatagggaga agggggggtg gttgtcggtc agcagcttga ttcacagcag
8401 gctcacaaga ttgcattctt tgaacagtag gctctagatg tctcagtaga taacctcaaa
8461 gagcatggcg ccagggagtg attgccctca gcgaaccttc tggcggctga agcagaagtg
8521 agtttgccca cattctgcat ttatgataaa cagtttgctg tttgatcata tagcctcagt
8581 ggaatgctga gttggtcaca acccatgggc attcagctct ctacatttcc ccctttctgt
8641 ttatgaatta attgaaagaa tgtaaggcca ggctgggcag ctctcatttt ctgattggca
8701 gtccatccaa ttttacagac tatgaacaga agacagagac aaaacaacat tattccaagg
8761 actacatata agatgttaat gtggtgcttt agacaggtcc aagggttgag gctctccagg
8821 ccttgctgga atttggtcta gtcttctaaa gaaggctgaa actcttgagt ttgtttattt
8881 aaatcaagaa ttttgttttg taattcaccg atatcaaagg tgatgttgga tgtgcaagct
8941 ccctgcaaat gggctttcac aaggtcccat ggatactcac tttggttata ttctaagttg
9001 gttacacaaa tatgagtgtg attaaaatga tggtgcaatt gctgttgcaa ttgcaagctt

9061 tgtacttgtt ctcctaactg tagagccgtg gatttcaaca atgccacttc agtttgtaac
9121 tcagtgttaa ttttattctg aagtggccat gcttggttgg ctatacgcat ccagttttcc
9181 atgtactgag ttgtttgaat agaactatgc aaagctacag aggacatcac aacggaagtt
9241 attaatgtga ccaaggaaac aatagcaaaa attatcatgc ctaaggctct atgggcacaa
9301 tgagtaagtt gagttagaag atgtttcatg aagtgcaaag caggtgtggt agcccaaggc
9361 tgggacagat taacaggaat ccatagccca gggatgcgac ccaaaattat caaagtagag
9421 acattatgtg tttgcaatgt gctatgatta atgcagtgac ataactggca agatttacag
9481 gtcaattgga tattgtttac ctggagctgg tccttcttag ttgccaaaaa agacataagg
9541 attaaaaaga caaaccataa a