

Nucleotide sequence of the cytochrome oxidase subunit I gene from rice mitochondria

Koh-ichi Kadowaki*, Takeshi Suzuki, Shigeru Kazama, Tepei Oh-fuchi and Wataru Sakamoto

Department of Molecular Biology, National Institute of Agrobiological Resources, Tsukuba 305, Japan
 Submitted August 11, 1989
 EMBL accession no. X15990

The cytochrome oxidase subunit I gene (*coxI*) from rice mitochondria was cloned and sequenced. Genomic library of rice mitochondrial DNA was constructed in Bluescript vector, and screened by using a 22 bases synthetic nucleotide encompassing the nucleotide region 673-694 of maize *coxI* (1). The amino acid sequence was deduced from the nucleotide sequence and translated according to the universal genetic code except that CGG was translated as tryptophan (W) as proposed (2). The predicted protein, 524 amino acids with a molecular weight of 57796 daltons, shows 99.2% homology with that of maize (1), 95.2% with soybean (3), 99.2% with sorghum (4), 93.3% with *Oenothera* (5), and 99.1% with wheat (6). The 3' untranslated region is highly AT-rich, together with a 29 bp direct repeat (boxed) and a possible stem-loop structure (arrowed).

```

          ggaattccaacacactctactgatctccaattagaagatcgaaagaggggggggaaataatgctgataccactt
-180 ttttttcaaagcaagcagatagtgacctaacaggactttgaaataactctgcaatttgaagacggagaaaagaaagctgagaagaat
-90 tatcgaaactggtttagatcaaggtctctctctccagcccccagccctctctgataagcttgaagatgttccaacagcaaaactcaa
  1 ATGACAAATCTGGTCCGATGGCTCTCCCACTAACCAAGGATATCGGTACTCTATTTTCACTTCCTGGGTGCCATGCAGGAGGTGATG
  M T N L V R W L F S T N H K D I G T L Y F I F G A I A G V M
  91 GGCACATGCTCTCCGACTGATTCGTATGGAAATAGCCCGACCCGGCGATCAAATCTGGTGGGAAATCATCACTTTATAATGTTTTTA
  G T C F S Y L I R M E L A R P G D Q I L G G N H G Q L Y N V L
  181 ATAACGGCTCAGCGCTTTTAAATGATCTTTTATGGTATGCGCGGAGATAGGTGGATTTGGTAAATGGTTTCTCCGATCTTGATA
  I T A H A F L M I F F M V H P A M I G F G N W F V P I L I
  271 GGTGCACCTGACATGGCAATTCACAGATAAATAATATATCATCTCGTGTGGCCACCAAGTCTCTGCTCCTAATAGCTCAGCCTTA
  G A P D M A F P R L N N I S F W L L P P S L L L L L S S A L
  361 GTAGAAGTGGCCAGCGCACGGTGGACGGTCTATCCGCCCTTAAGTGGTATTACCAGCCATTCTGGAGGAGCAGTTGATTTAGCAATT
  V E V G S G T G W T V Y P P L S G I T S H S G G A V D L A I
  451 TTTAGCTTTCATCTACAGGTGTTTCATCAATTTAGTTCATCAATTTTATAACAACACTATCTCAACATGCGTGGACCTGGAATGACT
  F S L H L S G V S S I L G S I N F I T T I F N M R G P G M T
  541 ATGCATAGATTACCACCTTTTGTGTGGTCCGTTCTAGTGACAGCATTCTCACTTTTATATCACTTCGGTACTGGCAGGGGCAATTACA
  M H R L P L F V W S V L V T A F L L L L S L P V L A G A I T
  631 ATGTTATTAACCGATCGAACTTTAATACAACCTTTTGTGATCTGACAGGAGGGAGACCCAAATATTATCCAGCATCTCTTTGGTTC
  H L L T D R N F N T T F F D P A G G D P I L Y Q H L P W F
  721 TTCGGTCAATCCAGAGGTATATCTCATCTCGCTGGATTCGGTATATTAATGATTCGTATCGACTTTTCAAGAAAACCGGCTCTTC
  F G H P E V Y A I L I L P G F G I I S H I V S F S R K V C F
  811 GGGTACTAGGCATGGTTATGCCATGATAAGTATAGGTGTTCTGGATTTCTAGTTGGGCTCATATATTTACTGTGGGCTTAGAC
  G Y L G M V Y A M I S I G V L G F L V W A H H M F T V W G L D
  901 GTTGATCGCGTGCCTACTCCGCGAGCTACCATGATCATAGCTGTGCCCACTGGAATAAAAATCTTTAGTTGGATGTCTACCATGTGG
  V D T R A Y F T A A T M I I A V P T G I K I F S W I A T M W
  991 GGAGGTTGATACAATAACAAACACCCATGTTATTGTGCTGATGGGTTTCATCTTTTGTTCACCATAGGAGGCTCACTGGAATGTTCTA
  G G S I Q Y K T P M L F A V G F I F L F T I G G L T G I V L
  1081 GCAAACTCTGGGCTAGACATGCTCTACATGATCTTATATGTGGTGGCATTTCCAATATGACTTCTACTTCTTCTAGGGAGCGGTTTGTCT
  A N S G L D I A L H D T Y Y V V A H F H Y V L S M G A V F A
  1171 TTATTGTCTGGATTTACTATTTGGGGGGTAAATCTTTGGTCCGACATATCTGAAACTTTAGGCCAAATCCATTTTGGATTAATCTTT
  L F A G F Y Y W V G R I F G W T Y P E T L G Q I H F W I T F
  1261 TTCGGGTTAATCTGACTCTTCTCCATGCAATCTTAGGCTTCGGGATGCCAGCTCCAGATTCCAGATTCCAGATTCCAGCTTACGCC
  F G V N L T F P M H F L G L S G H P R I P D Y P D Y A
  1351 GGAATGAACTCTGACGAGTTCGTTCTTATATATCCGTAGTTGGGATTCGCTGTTCTCGTAGTTGCGCAATCACTCAAGCAGT
  G W N A L S F G S Y I S V V G I R R F P V V V A I T S S S
  1441 GGAAAGACAAAAGATGTCCGGAAGCTCTTGGGCTGTGAACAGAAATCAACCACTAGAATGGTGGTCAAAAGCCCTCCAGCCCTT
  G K N K R C A E S P W A V E Q N P T L T E W L V Q S P P A P
  1531 CATACTTTGGAGAACTCCAGCTATCAAAGAGACAAAAGCTAGgcgcactctcctaactctgcccataaaaaataatattttaat
  H T F G E L P A I K E T K S *
  1621 atattattattatataatgcccctataaaaaataatattttaatatattattattctattctattgtgtgaa
          ▶◀
    
```

*To whom correspondence should be addressed.

References

- (1) Isaac, P.G. et al. (1985) EMBO J. 4, 1617-1623.
- (2) Fox, T.D. and Leaver, C.J. (1980) Cell 26, 315-323.
- (3) Grabau E.A. (1986) Plant Mol. Biol. 7, 377-384.
- (4) Bailey-Serres, J. et al. (1986) Cell 47, 567-576.
- (5) Hiesel, R. et al. (1987) EMBO J. 6, 29-34.
- (6) Bonen, L. et al. (1987) Nucleic Acids Res. 15, 6734.