

DNA Data Bank of Japan

DNA Database

Release 11, July 1992, including 65,693 entries, 84,839,075 bases.

This database may be copied and redistributed without permission, on the condition that all the statements in this release note are reproduced with each copy.

This release contains the newest data prepared by DDBJ, GenBank, and the EMBL Data Library as of July, 1992. Thanks to international collaboration between the three data banks, the International Nucleotide Sequence Database was organized, resulting in a unified database published in this release.

All the entries designated by the accession numbers with a "D" have been collected and processed by DDBJ, and the rest have been prepared by GenBank and the EMBL Data Library. It is particularly noted that for an unavoidable problem existing at present some entries in the three databases may overlap. We apologize for this, and are trying to resolve this problem by working together with the other data banks.

Because the release contains unified data, all the entries have been annotated with common feature keys.

The present release was prepared by processing the data on the relational database management system (Sybase). It does not include amino acid sequence data, because the genetic code system is known to be no longer uniform among species and organella, and we are not yet prepared for this.

Published by: K. Ikeo, Y. Ugawa, Y. Yamazaki, M. Horie, M. Iwase, M. Saito, Y. Sato, S. Suzuki, Y. Hasegawa, Y. Hattori, M. Hirashima, M. Shimoyama, E. Hatada, R. Suzuki, R. Uchida, Y. Shidahara, Y. Sakuma, A. Watanabe, Y. Ueda, E. Shimizu, T. Kawamoto, H. Kitakami, N. Saitou, T. Gojobori, and Y. Tateno

DNA Data Bank of Japan
National Institute of Genetics
Mishima 411, Japan
Phone: +81 559 75 0771
FAX: +81 559 75 6040
E-mail: ddbj@ddbj.nig.ac.jp

Acknowledgement: We are grateful to GenBank and the EMBL Data Library for permitting us to include their data in the present release.

This release covers 13 categories of organisms and others as follows:

ddbjbct.*** Category for bacteria
ddbjinv.*** Category for invertebrates
ddbjmam.*** Category for mammals
ddbjorg.*** Category for organella
ddbjphg.*** Category for phages
ddbjpln.*** Category for plants
ddbjpri.*** Category for primates
ddbjrna.*** Category for RNAs
ddbjrod.*** Category for rodents
ddbjsyn.*** Category for synthetic DNAs
ddbjuna.*** Category for unannotated sequences
ddbjvrl.*** Category for viruses
ddbjvrt.*** Category for vertebrates other than mammals, primates, and rodents

Each category then has the following eight files. Note that all the files except for *****.seq may include more than 80 characters in one line. If this is the case, the line is folded at every 81th column in the file on the distribution tape with the fixed record size of 80 bytes.

```
-----
*****.seq List of an entry in DDBJ format, see Table 1 for details
*****.acc List of the accession numbers, see Table 2 for details
*****.aut List of the authors, see Table 3 for details
*****.dir List of the short directory, see Table 4 for details
*****.idx List of indices, see Table 5 for details
*****.jou List of the journals, see Table 6 for details
*****.key List of the key words, see Table 7 for details
*****.org List of the species names, see Table 8 for details
-----
```

Table 1. Part of the contents in the file 'ddbjbct.seq'. This shows all pieces of information on one entry in DDBJ format.

```
-----
LOCUS      ABCAARAA      1624 bp ds-DNA      BCT      15-SEP-1990
DEFINITION A.aceti acetic acid resistance protein (aarA) gene, complete cds.
ACCESSION  M34830
KEYWORDS   acetic acid resistance protein.
SOURCE     A.aceti (strain 10-8) DNA, clone pAR1611.
  ORGANISM Acetobacter aceti
            Prokaryota; Bacteria; Gracilicutes; Scotobacteria;
            Aerobic rods and cocci; Azotobacteraceae.
REFERENCE  1 (bases 1 to 1624)
AUTHORS    Fukaya,M., Takemura,H., Okumura,H., Kawamura,Y., Horinouchi,S. and
            Beppu,T.
TITLE      Cloning of genes responsible for acetic acid resistance in
            acetobacter aceti
JOURNAL    J. Bacteriol. 172, 2096-2104 (1990)
STANDARD   simple staff_entry
FEATURES   Location/Qualifiers
  RBS      171..176
            /note="ribosome binding site (put.)"
  CDS      185..1495
            /note="acetic acid resistance protein (aarA)"
            /codon_start=1
  misc_signal 1508..1545
            /note="transcription termination signal"
BASE COUNT 400 a      446 c      404 g      374 t
ORIGIN
1 gcattgcattt gcacacattc gcgcgaccct aagcccaaaa aactgtgggt ttccaagcat
61 actcctttcc gataacgctt cgtttatcgc tggcaacctt ccggtttcct tttgaatgag
121 tgacaaagtg tgacgagcag gcgcgagcag cgaccgtggc ccaaccatgc agaaggaaac
181 actaatgagc gcgtcgcaga aagaaggtaa gctatctacc gctaccattt cggttgatgg
241 aaaatccgcc gaaatgcctg tgctttcagg cactctggga ccggatgta tgcacatccg
301 caaacttccg gcgcaactgg gcgttttcac gtttgacca ggttacgggg aaacagcggc
361 ctgcaacagc aaaatcacct ttattgatgg tgataaaggc gttctgctgc accgtggtta
421 ccctattgcg cagctggacg aaaatgcttc ctacgaagaa gttatttatc tgcttttgaa
481 tggcgaactg cccaacaagg tgcagtacga caccttcacc aacaccctta caaacatac
541 gctgctgcac gagcagatcc gtaacttctt taacggcttc cggcgtgatg cccaccaat
601 ggccattctg tgtggtacgg ttggggcctt gtctgccttc taccagatg ccaacgat
661 tgccattccc gccaatcggg atctggccgc catgcggctg attgccaaaa tccaacct
721 tgcggcattg gttacaat acacgcaggg tgaagcctt atctaccgc ggaatgatct
781 gaactacgca gaaaacttcc tgtccatgat gttcgcgcgc atgtccgaac cttacaaggt
841 caaccctggt ctggcccgcg ccatgaaccg gattctgatt ctgcatgccg atcatgagca
901 gaatgcctct acctccaccg tacgtctggc tggttctaca ggggccaatc cgtttgctg
961 tattgctgcg ggcattgccg ctctgtgggg acctgcacat ggtggcgcaa acgaactgt
1021 gctgaaaatg ctggcccgtg ttggaagaa agaaaatatt cctgccttta tcgacaggt
1081 gaaggacaag aacagcggcg taaagctgat gggctttggc caccgcgttt acaagaactt
1141 cgaccacgct gcgaagatca tgcagcagac ctgccacgaa gtgctgacag aactggcat
1201 taaggatgat ccgctgctgg atctggcggg tgagctggaa aagattgctc tgagcagatg
-----
```

```

1261 ttacttcgtg cagcgcaaac tttaccgaa tgtggatttc tactctggca tcattctcaa
1321 ggccatgggc atccccacca gtatgtttac tgtgctggtt gccgtagccc gcaccaccgg
1381 ctgggtgagc cagtggagg aaatgattga agaaccgggc cagcgtatca gccgccctcg
1441 ccagctttat attggcgcac cgcagcgtga ctatgtgccg cttgccaaac gctaaaacag
1501 actaacccaa aaagccgact tcccgtagg aaagtcggct tttgtttgc acgctgtttc
1561 caaaaaata gggcggcaga gcgaataaac gctacctage cttcaggcat aaaaaaacgc
1621 atgc

```

//

Table 2. Part of the contents in the file 'ddbjbct.acc'.

The first column refers to the secondary accession number, second column to the locus name, and third to the primary accession number. The primary number may be the same as the secondary number. They are arranged in the ascending order of the secondary accession numbers.

```

-----
D00001 -> ECOPBPAA    X04516
D00002 -> ECOPIYRH    X04469
D00006 -> PNS981TET    D00006
D00020 -> COLE2LYS    D00020
D00021 -> COLE31YS    D00021
D00038 -> BRLAM330    D00038
D00066 -> BAC139AC    D00066
D00067 -> ECONANA    M20207
D00069 -> ECOUVRD2    D00069
D00087 -> BACXYNAA    D00087
-----

```

Table 3. Part of the contents in the file 'ddbjbct.aut'.

For each author name given on the left to the arrow, the corresponding locus name and primary accession number are respectively listed on the right. They are arranged in the alphabetical order of the author names.

```

-----
Aan,F. -> STYCRR    X05210
Aan,F. -> STYENZI    M76176
Aaronson,W. -> ECOKPSD    M64977
Aaronson,W. -> ECONEUA    J05023
Abad-Lapuebla,M.A. -> VIBTDHI    D90238
Abdel-Mawgood,A.L. -> CYAPSBHA    X16394
Abdel-Meguid,S.S. -> TRNGDRECM    J01843
Abdelal,A. -> STYCARA    M36540
Abdelal,A. -> STYCARAB    X13200
Abdelal,A.H. -> PSENOSA    M60717
-----

```

Table 4. Part of the contents in the file 'ddbjbct.dir'.

For each locus name given in the first column, the corresponding primary accession number, molecular type, number of nucleotide pairs, and description for the locus are respectively listed. They are arranged in the alphabetical order of the locus names.

```

-----
ABCAARAA    M34830 ds-DNA    1624 A.aceti acetic acid resistance protein (aarA)
gene, complete cds.
ABCADHCC    D00635 ds-DNA    4230 A. polyoxogenes alcohol dehydrogenase (EC
1.1.99.8) and cytochrome c genes.
ABCALDH    D00521 ds-DNA    2683 A.polyoxogenes membrane-bound aldehyde
dehydrogenase gene, complete cds and flanks.
ABCBCSAA    M37202 ds-DNA    9540 A.xylinum bcs B, bcs C and bcs D genes,
complete cds and bcs A gene, partial cds.
ABCCELA    M76548 ds-DNA    1165 Acetobacter xylinum UDP pyrophosphorylase
(celA) gene, complete cds.
ABCCELSYN    X54676 ds-DNA    5363 A. xylinum gene for cellulose biosynthesis
ABCIS1380    D10043 ds-DNA    1665 A.pasteurianus insertion sequence IS1380.
-----

```

ACAADH1 D90004 ds-DNA 2467 Acetobacter aceti(K6033) alcohol
 dehydrogenase subunit gene(adh1).
 ACCAAC2 M62833 ds-DNA 1123 Acinetobacter baumannii aminoglycoside
 acetyltr ansferase (aac2) gene, complete cds.
 ACCACEAA M62822 ds-DNA 1874 A.baumannii chloramphenicol acetyltransferase
 (cat) gene, complete cds.

Table 5. Part of the contents in the file 'ddbjbct.idx'.
 The first column refers to the locus name, second column to the starting site
 of the locus in byte, and third to its ending site in byte. They are arranged
 in the alphabetical order of the locus names.

```

§*****
#ABCAARAA      0      3211
#ABCADHCC     3212     10608
#ABCALDH     10609     15864
#ABCBCSAA     15865     29583
#ABCCELA     29584     32289
#ABCCELSYN    32290     40960
#ABCIS1380    40961     44711
#ACAADH1     44712     49357
#ACCAAC2     49358     52395
  
```

Table 6. Part of the contents in the file 'ddbjbct.jou'.
 This gives information on the journal in which sequence data were published.

(in) Chaloupka,J. and Krumphanzl,V. (Eds.); Extracellular Enzymes of
 Microorganisms: 129-137, Plenum Press, New York (1987) -> BACAMYABS M57457
 (in) Ganesan,A.T., Chang,S. and Hoch,J.A. (Eds.); Molecular Cloning and Gene
 Regulation in Bacilli: 3-10, Academic Press, New York (1982) -> BACRG16S
 M55011
 (in) Ganesan,A.T., Chang,S. and Hoch,J.A. (Eds.); Molecular Cloning and Gene
 Regulation in Bacilli: 3-10, Academic Press, New York (1982) -> BACRG16SA
 M55006
 (in) Ganesan,A.T., Chang,S. and Hoch,J.A. (Eds.); Molecular Cloning and Gene
 Regulation in Bacilli: 3-10, Academic Press, New York (1982) -> BACRG16SB
 M55008
 (in) Hoch,J.A. and Setlow,P. (Eds.); Molecular Biology of Microbial
 Differentiation: 85-94, American Society for Microbiology, Washington, DC
 (1985) -> BACSPOII M57606
 (in) Holmgren,A. (Ed.); Thioredoxin and Glutaredoxin Systems: Structure and
 Function: 11-19, Unknown name, Unknown city (1986) -> ECOTRXA1 M54881
 (in) Kjeldgaard,N.C. and Maaloe,O. (Eds.); Control of ribosome synthesis:
 138-143, Academic Press, New York (1976) -> ECOLAC J01636
 (in) Losick,R. and Chamberlin,M. (Eds.); RNA polymerase: 455-472, Cold Spring
 Harbor Laboratory, Cold Spring Harbor, NY (1976) -> ECOTGY1 K01197
 (in) Sikes,C.S. and Wheeler,A.P. (Eds.); Surface reactive peptides and
 polymers. Discovery and commercialization.: 186-200, American Chemical
 Society, Washington, D.C. (1991) -> ECOTGP J01714
 (in) Sund,H. and Blauer,G. (Eds.); Protein-Ligand Interactions: 193-207,
 Walter de Gruyter, New York (1975) -> ECOLAC J01636
 (in) Wu,R. and Grossman,L. (Eds.); Methods in Enzymology, Recombinant DNA,
 part E: In press, Academic Press, New York, N.Y. (1986) -> PLMCG M11320
 Acta Microbiol. Pol. 35, 175-190 (1986) -> ECOTGG1 M54893
 Actinomycetologica 5, 14-17 (1991) -> STMARGG D00799
 Adv. Biophys. 21, 115-133 (1986) -> R10REP M26840
 Adv. Biophys. 21, 175-192 (1986) -> ECONUSAA M26839
 Adv. Enzyme Regul. 21, 225-237 (1983) -> ECOPURFA M26893
 Adv. Exp. Med. Biol. 195, 239-246 (1986) -> ECOAPT M14040
 Agric. Biol. Chem. 50, 2155-2158 (1986) -> ECONANA M20207
 Agric. Biol. Chem. 50, 2771-2778 (1986) -> BRLAM330 D00038
 Agric. Biol. Chem. 51, 2019-2022 (1987) -> BACCGT D00129

Agric. Biol. Chem. 51, 2641-2648 (1987) ->	STRSAGP	D00219
Agric. Biol. Chem. 51, 2807-2809 (1987) ->	BACPGECR	M35503
Agric. Biol. Chem. 51, 3133-3135 (1987) ->	BACXYLAP	D00312
Agric. Biol. Chem. 51, 455-463 (1987) ->	BACHDCRY	D00117
Agric. Biol. Chem. 51, 953-955 (1987) ->	BACXYNAA	D00087
Agric. Biol. Chem. 52, 1565-1573 (1988) ->	BACIP135	D00348
Agric. Biol. Chem. 52, 1785-1789 (1988) ->	BACTMR	D00343
Agric. Biol. Chem. 52, 2243-2246 (1988) ->	PSEGI	D00342
Agric. Biol. Chem. 52, 399-406 (1988) ->	BACAMYEB	M35517
Agric. Biol. Chem. 52, 479-487 (1988) ->	ECAPALI	D00217

Table 7. Part of the contents in the file 'ddbjbct.key'.
For the locus and accession number respectively given on the right to the
arrow, the corresponding key words are listed on the left.

A.aceti acetic acid resistance protein (aarA) gene, complete cds. ->
ABCAARAA M34830
acetic acid resistance protein. -> ABCAARAA M34830
Cloning of genes responsible for acetic acid resistance in acetobacter aceti
-> ABCAARAA M34830
A. polyoxogenes alcohol dehydrogenase (EC 1.1.99.8) and cytochrome c genes.
-> ABCADHCC D00635
alcohol dehydrogenase; cytochrome c. -> ABCADHCC D00635
Cloning and sequencing of the gene cluster encoding two subunits of
membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes ->
ABCADHCC D00635
These data kindly submitted in computer readable form by: Toshimi Tamaki
Nakano Central Biochemical Institute 2-6 Nakamura-cho Handa-shi, Aichi-ken 475
Japan Phone: 0569-21-3331 Fax: 0569-23-8486 -> ABCADHCC D00635
A.polyoxogenes membrane-bound aldehyde dehydrogenase gene, complete cds and
flanks. -> ABCALDH D00521
aldehyde dehydrogenase gene; ethanol oxidation; membrane-bound enzyme. ->
ABCALDH D00521
Nucleotide sequence of the membrane-bound aldehyde dehydrogenase gene from
Acetobacter polyoxogenes -> ABCALDH D00521

Table 8. Part of the contents in the file 'ddbjbct.org'.
For the locus and accession number given on the right to the arrow, the
corresponding taxonomic names are respectively listed on the left. They are
arranged in the alphabetical order of the species names.

A. nidulans 6301 DNA. Anacystis nidulans Prokaryota; Bacteria; Gracilicutes;
Oxyphotobacteria; Cyanobacteria. -> ANIRUBPS X00019
A. nidulans DNA, clone pAN4. Anacystis nidulans Prokaryota; Bacteria;
Gracilicutes; Oxyphotobacteria; Cyanobacteria. -> ANIRGGX X00343
A. nidulans DNA. Anacystis nidulans Prokaryota; Bacteria; Gracilicutes;
Oxyphotobacteria; Cyanobacteria. -> ANIRGG X00512
A. polyoxogenes genomic DNA. Acetobacter polyoxogenes Prokaryota; Bacteria;
Gracilicutes; Scotobacteria; Aerobic rods and cocci; Azotobacteraceae. ->
ABCADHCC D00635
A. quadruplicatum (strain PR-6) DNA, clone pAQPR1. Agmenellum quadruplicatum
Prokaryota; Bacteria; Gracilicutes; Oxyphotobacteria; Cyanobacteria. ->
AQUPCAB K02660
A. quadruplicatum (strain PR6) DNA. Agmenellum quadruplicatum Prokaryota;
Bacteria; Gracilicutes; Oxyphotobacteria; Cyanobacteria. -> AQUCPCAB
K02659
A. vinelandii DNA. Azotobacter vinelandii Prokaryota; Bacteria; Gracilicutes;
Scotobacteria; Aerobic rods and cocci; Azotobacteraceae. -> AVINIFUSV
M17349
A.aceti (strain 10-8) DNA, clone pAR1611. Acetobacter aceti Prokaryota;
Bacteria; Gracilicutes; Scotobacteria; Aerobic rods and cocci;
Azotobacteraceae. -> ABCAARAA M34830

A.actinomycetemcomitans (strain JP2) DNA, clone lambda-OP8. Actinobacillus
actinomycetemcomitans Prokaryota; Bacteria; Gracilicutes; Scotobacteria;
Facultatively anaerobic rods; Pasteurellaceae. -> ACNLKTXN M27399
A.anitratum DNA, clone pLJD1. Acinetobacter anitratum Prokaryota; Bacteria;
Gracilicutes; Scotobacteria; Neisseriaceae. -> ACCCITSYN M33037

The files in this release are arranged in the following order with non-labeled
format.

Release note

FILE.001 ddbjrel.txt
Category for bacteria, 7269 entries, 12576601 bases
FILE.002 ddbjbct.acc
FILE.003 ddbjbct.aut
FILE.004 ddbjbct.dir
FILE.005 ddbjbct.idx
FILE.006 ddbjbct.jou
FILE.007 ddbjbct.key
FILE.008 ddbjbct.org
FILE.009 ddbjbct.seq
Category for invertebrates, 5250 entries, 7462407 bases
FILE.010 ddbjinv.acc
FILE.011 ddbjinv.aut
FILE.012 ddbjinv.dir
FILE.013 ddbjinv.idx
FILE.014 ddbjinv.jou
FILE.015 ddbjinv.key
FILE.016 ddbjinv.org
FILE.017 ddbjinv.seq
Category for mammals, 2459 entries, 3208845 bases
FILE.018 ddbjmam.acc
FILE.019 ddbjmam.aut
FILE.020 ddbjmam.dir
FILE.021 ddbjmam.idx
FILE.022 ddbjmam.jou
FILE.023 ddbjmam.key
FILE.024 ddbjmam.org
FILE.025 ddbjmam.seq
Category for organella, 2115 entries, 3314738 bases
FILE.026 ddbjorg.acc
FILE.027 ddbjorg.aut
FILE.028 ddbjorg.dir
FILE.029 ddbjorg.idx
FILE.030 ddbjorg.jou
FILE.031 ddbjorg.key
FILE.032 ddbjorg.org
FILE.033 ddbjorg.seq
Category for phages, 761 entries, 1116646 bases
FILE.034 ddbjphg.acc
FILE.035 ddbjphg.aut
FILE.036 ddbjphg.dir
FILE.037 ddbjphg.idx
FILE.038 ddbjphg.jou
FILE.039 ddbjphg.key
FILE.040 ddbjphg.org
FILE.041 ddbjphg.seq
Category for plants, 5571 entries, 9398767 bases
FILE.042 ddbjpln.acc
FILE.043 ddbjpln.aut
FILE.044 ddbjpln.dir
FILE.045 ddbjpln.idx
FILE.046 ddbjpln.jou
FILE.047 ddbjpln.key
FILE.048 ddbjpln.org

FILE.049 ddbjpln.seq
Category for primates, 14812 entries, 15742165 bases
FILE.050 ddbjpri.acc
FILE.051 ddbjpri.aut
FILE.052 ddbjpri.dir
FILE.053 ddbjpri.idx
FILE.054 ddbjpri.jou
FILE.055 ddbjpri.key
FILE.056 ddbjpri.org
FILE.057 ddbjpri.seq
Category for RNAs, 2570 entries, 1388993 bases
FILE.058 ddbjrna.acc
FILE.059 ddbjrna.aut
FILE.060 ddbjrna.dir
FILE.061 ddbjrna.idx
FILE.062 ddbjrna.jou
FILE.063 ddbjrna.key
FILE.064 ddbjrna.org
FILE.065 ddbjrna.seq
Category for rodents, 11887 entries, 13069117 bases
FILE.066 ddbjrod.acc
FILE.067 ddbjrod.aut
FILE.068 ddbjrod.dir
FILE.069 ddbjrod.idx
FILE.070 ddbjrod.jou
FILE.071 ddbjrod.key
FILE.072 ddbjrod.org
FILE.073 ddbjrod.seq
Category for synthetic DNAs, 1258 entries, 821051 bases
FILE.074 ddbjsyn.acc
FILE.075 ddbjsyn.aut
FILE.076 ddbjsyn.dir
FILE.077 ddbjsyn.idx
FILE.078 ddbjsyn.jou
FILE.079 ddbjsyn.key
FILE.080 ddbjsyn.org
FILE.081 ddbjsyn.seq
Category for unannotated sequences, 2009 entries, 1829147 bases
FILE.082 ddbjuna.acc
FILE.083 ddbjuna.aut
FILE.084 ddbjuna.dir
FILE.085 ddbjuna.idx
FILE.086 ddbjuna.jou
FILE.087 ddbjuna.key
FILE.088 ddbjuna.org
FILE.089 ddbjuna.seq
Category for viruses, 6748 entries, 11329780 bases
FILE.090 ddbjvrl.acc
FILE.091 ddbjvrl.aut
FILE.092 ddbjvrl.dir
FILE.093 ddbjvrl.idx
FILE.094 ddbjvrl.jou
FILE.095 ddbjvrl.key
FILE.096 ddbjvrl.org
FILE.097 ddbjvrl.seq
Category for vertebrates other than mammals, primates, and rodents, 2984
entries, 3580818 bases
FILE.098 ddbjvrt.acc
FILE.099 ddbjvrt.aut
FILE.100 ddbjvrt.dir
FILE.101 ddbjvrt.idx
FILE.102 ddbjvrt.jou
FILE.103 ddbjvrt.key
FILE.104 ddbjvrt.org
FILE.105 ddbjvrt.seq
DNA data submission form

FILE.106 ddbjsub.txt