DNA Data Bank of Japan

DNA Database

Release 12, January 1993, including 97,683 entries, 120,815,244 bases.

The present release contains the newest data prepared by DDBJ, GenBank/NCBI, and the EMBL Data Library as of January, 1993. 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.

This release was prepared by processing the data on the relational data base management system (Sybase). Because the release contains unified data, all the entries have been annotated with common feature keys.

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/NCBI and the EMBL Data Library. 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.

Note in particular that the present release includes duplicated entries over the data files. This was originally caused by the fact that GenBank/NCBI no longer has a separate file for organella, and data belonging to this category were allocated to the other files according to the "host" species. For example, a human mitochondrial DNA sequence now belongs to the primate file in the GenBank/NCBI database. We, however, still maintain the organella file in this release. To revive the organella file in the GenBank/NCBI database and incorporate it in this release was nevertheless quite cumbersome and time-consuming, and did not allow us to get complete results. Namely, the pertinent files other than the organella file still include the data from organella. Thus please be careful when you want to get some results only for nuclear sequences.

From this release we also included EST(Expressed Sequence Tag)/TSF(Transcribed Sequence Fragment) data. Though we created a separate file for this category, we could not completely separate it from the other files, and could not but let the EST/TSF file fall in the same situation as in the organella file. That is, there are duplicated entries similar to those of organella data. Thus again be careful.

The present release 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: Y. Ugawa, Y. Yamazaki, K. Ikeo, M. Horie, M. Iwase, M. Saito, Y. Sato, S. Suzuki, Y. Hasegawa, Y. Hattori, M. Hirashima, M. Shimoyama, A. Hasegawa, H. Tsutsui, E. Hatada, R. Suzuki, R. Uchida, Y. Shidahara, M. Gojobori, Y. Sakuma, A. Watanabe, Y. Ueda, 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
ddbjsub@ddbj.nig.ac.jp (for data submissions)

ddbjupdt@ddbj.nig.ac.jp (for updates and notification of publication)

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

This release covers 14 categories of organisms and others as follows:

```
ddbjbct.*** Category for bacteria
ddbjest.*** Category for EST (expressed sequence tag)
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 nine files. Note that all the files except for ddbj***.seq and ddbj***.sdr 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.

```
ddbj***.seq  List of an entry in DDBJ format, see Table 1.
ddbj***.acc  List of the accession numbers, see Table 2 .
ddbj***.aut  List of the authors, see Table 3.
ddbj***.dir  List of the short directory in DDBJ style, see Table 4.
ddbj***.idx  List of indices, see Table 5.
ddbj***.jou  List of the journals, see Table 6.
ddbj***.key  List of the key words, see Table 7.
ddbj***.org  List of the species names, see Table 8.
ddbj***.sdr  List of the short directory in GenBank style, see Table 9.
```

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 gcatgcattt gcacacattc gcgcgaccct aagcccaaaa aactgtggtt ttccaagcat
       61 actcctttcc gataacgctt cgtttatcgc tggcaacctt ccggtttcct tttgaatgag
      121 tgacaaagtg tgacgagcag gccgcagcag cgaccgtggc ccaaccatgc agaaggaaac
      181 actaatgagc gcgtcgcaga aagaaggtaa gctatctacc gctaccattt cggttgatgg
      241 aaaatccgcc gaaatgcctg tgctttcagg cactctggga ccggatgtta tcgacatccg
      301 caaacttccg gcgcaactgg gcgttttcac gtttgaccca ggttacgggg aaacagcggc
      361 ctgcaacagc aaaatcacct ttattgatgg tgataaaggc gttctgctgc accgtggtta
      421 ccctattgcg cagctggacg aaaatgcttc ctacgaagaa gttatttatc tgcttttgaa
      481 tggcgaactg cccaacaagg tgcagtacga caccttcacc aacaccctta caaaccatac
      541 gctgctgcac gagcagatcc gtaacttctt taacggcttc cggcgtgatg cccacccaat
      601 ggccattctg tgtggtacgg ttggggcttt gtctgccttc tacccagatg ccaacgatat
      661 tgccattccc gccaatcggg atctggccgc catgcggctg attgccaaaa tcccaaccat
      721 tgcggcatgg gcttacaaat acacgcaggg tgaagccttt atctacccgc ggaatgatct
      781 gaactacgca gaaaacttcc tgtccatgat gttcgcgcgc atgtccgaac cttacaaggt
      841 caaccctgtt ctggcccgcg ccatgaaccg gattctgatt ctgcatgccg atcatgagca
      901 gaatgcctct acctccaccg tacgtctggc tggttctaca ggggccaatc cgtttgcctg
      961 tattgctgcg ggcattgccg ctctgtgggg acctgcacat ggtggcgcaa acgaagctgt
     1021 gctgaaaatg ctggcccgta ttggcaagaa agaaaatatt cctgccttta tcgcacaggt
     1081 gaaggacaag aacagcggcg taaagctgat gggctttggc caccgcgttt acaagaactt
     1141 cgacccacgt gcgaagatca tgcagcagac ctgccacgaa gtgctgacag aacttggcat
     1201 taaggatgat ccgctgctgg atctggcggt tgagctggaa aagattgctc tgagcgatga
     1261 ttacttcgtg cagcgcaaac tttacccgaa tgtggatttc tactctggca tcattctcaa
     1321 ggccatgggc atccccacca gtatgtttac tgtgctgttt gccgtagccc gcaccaccgg
     1381 ctgggtgagc cagtggaagg aaatgattga agaaccgggc cagcgtatca gccgccctcg
     1441 ccagctttat attggcgcac cgcagcgtga ctatgtgccg cttgccaaac gctaaaacag
     1501 actaacccaa aaagccgact tcccgtaagg aaagtcggct ttttgtttgc acgctgtttc
     1561 caaaaaaata gggcggcaga gcgaataaac gctacctagc 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 -> ECOPYRH    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 short directory in DDBJ style 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 acetyltransferase (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 respectively given on the right to the arrow, the corresponding taxonomic names are 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

Table 9. Part of the short directory file in GenBank style in the file 'ddbjbct.sdr'.
The short directory file contains brief descriptions of all of the sequence entries contained in the GenBank style.

ABCAARAA	A.aceti acetic acid resistance protein (aarA) gene, complete	1624bp
ABCADHCC	A. polyoxogenes alcohol dehydrogenase (EC 1.1.99.8) and	4230bp
ABCALDH	A.polyoxogenes membrane-bound aldehyde dehydrogenase gene,	2683bp
ABCBCSABCD	A.xylinum bcs A, B, C and D genes, complete cds's.	9540bp
ABCCELA	Acetobacter xylinum UDP pyrophosphorylase (celA) gene,	1165bp
ABCCELSYN	A. xylinum gene for cellulose biosynthesis	5363bp
ABCIS1380	A.pasteurianus insertion sequence IS1380.	1665bp
ACAADH1	Acetobacter aceti(K6033) alcohol dehydrogenase subunit	2467bp
ACCAAC2	Acinetobacter baumannii aminoglycoside acetyltransferase	1123bp
ACCACEAA	A.baumannii chloramphenicol acetyltransferase (cat) gene,	1874bp
ACCAPHA6	Acinetobacter baumannii aphA-6 gene.	1170bp
ACCBENABCA	A.calcoaceticus BenA, BenB, BenC, BenD, and BenE proteins	15922bp
ACCCAT	Acinetobacter calcoaceticus cat operon.	15922bp
ACCCATAM	A.calcoaceticus catA and catM genes, encoding catechol 1,	5537bp
ACCCHMO	Acinetobacter sp. cyclohexanone monooxygenase gene, complete	2128bp
ACCCITSYN	A.anitratum citrate synthase gene, complete cds.	1895bp

---

In addition to the 9 tables the five following index files are included in this release. These files were prepared irrespective of the 14 categories of taxonomic divisions.

- Accession number index file
- Keyword phrase index file
- Author name index file
- Journal citation index file
- Gene name index file

A brief description is given for each file in the following.

Table 10. Part of the accession number index file in the 'ddbjacc.idx'.
The following excerpt from the accession number index file illustrates the format of the index. Note that as mentioned above there are such a case where an accession number for a taxonomic category is the same as that for EST or ORG; for example, PRI D12345 and EST D12345 under the same accession number D12345.

M33790	SHFINVEA	BCT M33790
M33791	BACORF2	BCT M33791

```
M33792        FTRCPRBCLC ORG X55829 FTRCPRBCLC PLN X55829
M33793        FTRCPPRBCL ORG X55830 FTRCPPRBCL PLN X55830
M33794        ATPCPARRBC ORG X55831 ATPCPARRBC PLN X55831 ATPCPRBCLB ORG X15925
              ATPCPRBCLB PLN X15925
M33796        NRACPNTRBC ORG X55827 NRACPNTRBC PLN X55827
M33797        NRACPRBCL  ORG X55828 NRACPRBCL  PLN X55828
M33798        ACCPCACGH  BCT M33798
M33799        PSETRPEGDC BCT M33799
```

Table 11. Part of the keyword phrase index file in the 'ddbjkey.idx'.
Keyword phrases consist of names for gene products and other characteristics of sequence entries.

```
A CHANNEL
              DROCHA      INV M17155
A COMPONENT
              SQLCVEA     VRL M38183
A LOCUS
              GORGOGOA3   PRI X54375 GORGOGOA4   PRI X54376
A LOCUS ALLELE
              GORA0101    PRI X60258 GORA0201    PRI X60259 GORA0401    PRI X60257
              GORA0501    PRI X60256
A MULTI-GENE FAMILY
              RICGLUTE    PLN D00584
A PROTEIN
              MS2AAR      PHG M25187 ST1APCS     PHG M25396
A SEQUENCE
              HS5TOA30    VRL D00148 HS5TOA31    VRL D00147
```

Table 12. Part of the author name index file in 'ddbjaut.idx'.
The author name index file lists all of the author names that appear in the citations.

```
ABE,A.
              HUMMHDRBWE PRI M27509 HUMMHDRBWF PRI M27510 HUMMHDRBWG PRI M27511
              YSCGAL11A  PLN M22481
ABE,C.
              S85445     BCT S85445
ABE,E.
              M23442     UNA M23442
ABE,H.
              CHKADF     VRT M55660 CHKCOF     VRT M55659
ABE,K.
              CHPCLAC    PRI D11383 CHPIMRF    PRI D11384 CUGCUR09   PLN X64110
              CUGCUR37   PLN X64111 HPCCEXPA   VRL M55970 HPCCPEP1   VRL D10687
              HPCCPEP2   VRL D10688 HPCHABC82  VRL X51587 HPCNS2APA  VRL M55972
              HPCNS2PA   VRL M55971 HPCNS2PB   VRL M55973 HPCNS5PA   VRL M55974
              MUSKE2     ROD M65255 MUSKE2A    ROD M65256 MZECYS     PLN D10622
              RICCPI     PLN J03469 RICGLUTE   PLN D00584 RICLNOCI   PLN J05595
              RICOCS     PLN M29259 RICORYII   PLN X57658 RICOZA     PLN D90406
              RICOZB     PLN D90407 RICOZC     PLN D90408 S54524     PLN S54524
              S54526     PLN S54526 S54530     PLN S54530 S73960     ROD S73960
```

Table 13. Part of the journal citation index file in 'ddbjjou.idx'.
The journal citation index file lists all of the citations that appear in the references.

```
ACTA BIOCHIM. BIOPHYS. SIN. 23, 246-253 (1992)
              HUMPLASINS PRI M98056
ACTA BIOCHIM. POL. 24, 301-318 (1977)
              LUPTRFJ    RNA K00345 LUPTRFN    RNA K00346
ACTA BIOCHIM. POL. 26, 369-381 (1979)
              BLYTRNPHE  PLN X02683
```

```
ACTA BIOCHIM. POL. 29, 143-149 (1982)
             EMEMTA     ORG M32572 EMEMTA     PLN M32572 EMEMTB     ORG M32573
             EMEMTB     PLN M32573 EMEMTC     ORG M32574 EMEMTC     PLN M32574
             EMEMTD     ORG M32575 EMEMTD     PLN M32575 EMEMTE     ORG M32576
             EMEMTE     PLN M32576
ACTA BIOCHIM. POL. 34, 21-27 (1987)
             LUPNOSP    PLN M32571
--------------------------------------------------------------------------------

Table 14. Part of the gene name index file in 'ddbjgen.idx'.
This file lists all the gene names that appear in the feature table.
--------------------------------------------------------------------------------
AACC8
             STMAACC8   BCT M55426
AACC9
             MPUAACC9   BCT M55427
AACT
             HUMA1ACM   PRI K01500 HUMA1ACMA  PRI X00947 HUMA1ACMB  PRI M18035
             HUMAACT1   PRI M18906 HUMAACT2   PRI M22533 HUMAACTA   PRI J05176
AAD
             INTINTORF  BCT L06418 LMOMO229D  BCT X17478
AAD A1
             ENTAAC3VI  BCT M88012
AAD9
             ENEAAD9A   BCT M69221
AADA
             LMOMO229A  BCT X17479 S52249     BCT S52249 SYNAADA    SYN M60473
             TRNTAAB    BCT M55547 TRNTN21CAS BCT M86913
--------------------------------------------------------------------------------


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

Release note
    FILE.001 ddbjrel.txt          651 records
Category for bacteria, 10872 entries, 19000740 bases
    FILE.002 ddbjbct.acc        12923 records
    FILE.003 ddbjbct.aut        43054 records
    FILE.004 ddbjbct.dir        10872 records
    FILE.005 ddbjbct.idx        10873 records
    FILE.006 ddbjbct.jou        13745 records
    FILE.007 ddbjbct.key        41073 records
    FILE.008 ddbjbct.org        10872 records
    FILE.009 ddbjbct.sdr        10872 records
    FILE.010 ddbjbct.seq       679969 records
Category for EST (expressed sequence tag), 7346 entries, 2420026 bases
    FILE.011 ddbjest.acc         7347 records
    FILE.012 ddbjest.aut        67226 records
    FILE.013 ddbjest.dir         7346 records
    FILE.014 ddbjest.idx         7347 records
    FILE.015 ddbjest.jou         7359 records
    FILE.016 ddbjest.key        24942 records
    FILE.017 ddbjest.org         7346 records
    FILE.018 ddbjest.sdr         7346 records
    FILE.019 ddbjest.seq       199208 records
Category for invertebrates, 7779 entries, 11492617 bases
    FILE.020 ddbjinv.acc         9024 records
    FILE.021 ddbjinv.aut        31555 records
    FILE.022 ddbjinv.dir         7779 records
    FILE.023 ddbjinv.idx         7780 records
    FILE.024 ddbjinv.jou         9444 records
    FILE.025 ddbjinv.key        28679 records
    FILE.026 ddbjinv.org         7779 records
    FILE.027 ddbjinv.sdr         7779 records
    FILE.028 ddbjinv.seq       441401 records
```

```
Category for mammals, 3625 entries, 4713662 bases
    FILE.029 ddbjmam.acc        4356 records
    FILE.030 ddbjmam.aut       16725 records
    FILE.031 ddbjmam.dir        3625 records
    FILE.032 ddbjmam.idx        3626 records
    FILE.033 ddbjmam.jou        4240 records
    FILE.034 ddbjmam.key       13374 records
    FILE.035 ddbjmam.org        3625 records
    FILE.036 ddbjmam.sdr        3625 records
    FILE.037 ddbjmam.seq      194328 records
Category for organella, 2737 entries, 4368408 bases
    FILE.038 ddbjorg.acc        2979 records
    FILE.039 ddbjorg.aut        9763 records
    FILE.040 ddbjorg.dir        2737 records
    FILE.041 ddbjorg.idx        2738 records
    FILE.042 ddbjorg.jou        3459 records
    FILE.043 ddbjorg.key       10005 records
    FILE.044 ddbjorg.org        2737 records
    FILE.045 ddbjorg.sdr        2737 records
    FILE.046 ddbjorg.seq      163539 records
Category for phages, 885 entries, 1271991 bases
    FILE.047 ddbjphg.acc        1056 records
    FILE.048 ddbjphg.aut        3277 records
    FILE.049 ddbjphg.dir         885 records
    FILE.050 ddbjphg.idx         886 records
    FILE.051 ddbjphg.jou        1225 records
    FILE.052 ddbjphg.key        3402 records
    FILE.053 ddbjphg.org         885 records
    FILE.054 ddbjphg.sdr         885 records
    FILE.055 ddbjphg.seq       50903 records
Category for plants, 10740 entries, 18382781 bases
    FILE.056 ddbjpln.acc       12299 records
    FILE.057 ddbjpln.aut       41261 records
    FILE.058 ddbjpln.dir       10740 records
    FILE.059 ddbjpln.idx       10741 records
    FILE.060 ddbjpln.jou       13033 records
    FILE.061 ddbjpln.key       40185 records
    FILE.062 ddbjpln.org       10740 records
    FILE.063 ddbjpln.sdr       10740 records
    FILE.064 ddbjpln.seq      658620 records
Category for primates, 20892 entries, 21904974 bases
    FILE.065 ddbjpri.acc       24787 records
    FILE.066 ddbjpri.aut      116474 records
    FILE.067 ddbjpri.dir       20892 records
    FILE.068 ddbjpri.idx       20893 records
    FILE.069 ddbjpri.jou       25146 records
    FILE.070 ddbjpri.key       77017 records
    FILE.071 ddbjpri.org       20892 records
    FILE.072 ddbjpri.sdr       20892 records
    FILE.073 ddbjpri.seq     1023230 records
Category for RNAs, 3092 entries, 1736277 bases
    FILE.074 ddbjrna.acc        3416 records
    FILE.075 ddbjrna.aut       12319 records
    FILE.076 ddbjrna.dir        3092 records
    FILE.077 ddbjrna.idx        3093 records
    FILE.078 ddbjrna.jou        3470 records
    FILE.079 ddbjrna.key       10635 records
    FILE.080 ddbjrna.org        3092 records
    FILE.081 ddbjrna.sdr        3092 records
    FILE.082 ddbjrna.seq      111281 records
Category for rodents, 16031 entries, 17919502 bases
    FILE.083 ddbjrod.acc       19206 records
    FILE.084 ddbjrod.aut       72014 records
    FILE.085 ddbjrod.dir       16031 records
    FILE.086 ddbjrod.idx       16032 records
    FILE.087 ddbjrod.jou       19116 records
```

```
    FILE.088 ddbjrod.key       58899 records
    FILE.089 ddbjrod.org       16031 records
    FILE.090 ddbjrod.sdr       16031 records
    FILE.091 ddbjrod.seq      804025 records
Category for synthetic DNAs, 1361 entries, 1108671 bases
    FILE.092 ddbjsyn.acc        1474 records
    FILE.093 ddbjsyn.aut        5773 records
    FILE.094 ddbjsyn.dir        1361 records
    FILE.095 ddbjsyn.idx        1362 records
    FILE.096 ddbjsyn.jou        1552 records
    FILE.097 ddbjsyn.key        4765 records
    FILE.098 ddbjsyn.org        1361 records
    FILE.099 ddbjsyn.sdr        1361 records
    FILE.100 ddbjsyn.seq       55348 records
Category for unannotated sequences, 1661 entries, 1525342 bases
    FILE.101 ddbjuna.acc        1726 records
    FILE.102 ddbjuna.aut        6727 records
    FILE.103 ddbjuna.dir        1661 records
    FILE.104 ddbjuna.idx        1662 records
    FILE.105 ddbjuna.jou        1665 records
    FILE.106 ddbjuna.key        5047 records
    FILE.107 ddbjuna.org        1661 records
    FILE.108 ddbjuna.sdr        1661 records
    FILE.109 ddbjuna.seq       55043 records
Category for viruses, 9833 entries, 14691864 bases
    FILE.110 ddbjvrl.acc       11346 records
    FILE.111 ddbjvrl.aut       54586 records
    FILE.112 ddbjvrl.dir        9833 records
    FILE.113 ddbjvrl.idx        9834 records
    FILE.114 ddbjvrl.jou       12456 records
    FILE.115 ddbjvrl.key       36497 records
    FILE.116 ddbjvrl.org        9833 records
    FILE.117 ddbjvrl.sdr        9833 records
    FILE.118 ddbjvrl.seq      551459 records
Category for vertebrates other than mammals, primates, and rodents, 4486
entries, 5261256 bases
    FILE.119 ddbjvrt.acc        5174 records
    FILE.120 ddbjvrt.aut       18560 records
    FILE.121 ddbjvrt.dir        4486 records
    FILE.122 ddbjvrt.idx        4487 records
    FILE.123 ddbjvrt.jou        5298 records
    FILE.124 ddbjvrt.key       16329 records
    FILE.125 ddbjvrt.org        4486 records
    FILE.126 ddbjvrt.sdr        4486 records
    FILE.127 ddbjvrt.seq      227004 records
DNA data submission form
    FILE.128 ddbjsub.txt         336 records
Accession number index file
    FILE.129 ddbjacc.idx      105551 records
Keyword phrase index file
    FILE.130 ddbjkey.idx      101753 records
Author name index file
    FILE.131 ddbjaut.idx      268583 records
Journal citation index file
    FILE.132 ddbjjou.idx      100323 records
Gene name index file
    FILE.133 ddbjgen.idx       32193 records
```