

核酸塩基配列データベース

医療技術短期大学部 中島 広志

金沢大学計算機に登録、公開されている核酸塩基配列データの内容と利用法について述べる。

このデータは、京都大学化学研究所 大井 龍夫教授の御厚意により提供されたデータで Gen Bank (米国 Los Alamos 研究所) のデータを原データとしている。1984年3月現在約260万塩基が収録されており、374トラックに相当する。遺伝子データである核酸塩基配列データの説明や利用方法について、すでに多くの報告がなされている(1-5)。

ここでは、計算機によるこのデータベースの利用法及びデータの入出力のプログラムについて述べる。

計算機の利用については、コマンドの使用例をあげ、簡単な説明をおこなった。

なお、当計算機センターには、わかりやすいマニュアルが数多く整備されているのでコマンドの詳しい説明はこれらマニュアルを参照していただきたい。

このデータベースの使用料は無料で使用に際して何の条件もありませんから、自由に利用して下さい。遺伝情報が簡単に手にはいり、遺伝子に対する親近感が増すものと思います。

1. 核酸塩基配列データの内容

このデータは区分データセット CC0892. KAKUSAN. DATAにGENE1、GENE2、…GENE10と10個のメンバー名で登録されている。表1に10個のメンバーに含まれている内容を示す。データの見本として CC0892. KAKUSAN. DATA (GENE4)の一部を図1に示す。

図1において行番号10~450が1つのentry、又行番号460~640が1つのentryとなっている。つまり個々のデータはMemberから始まり、//で終わる一定のFORMATのもとに書かれている。またデータは72カラム内に書かれている。

ところで、原データの変更を防ぐ目的で、データセット CC0892. KAKUSAN. DATAには書き込み禁止のプロテクトをつけてある。このため、ユーザーはEDITコマンドによる END_SAVE又はPFDモードでのEDITによるデータの変更ができません。この不都合は、CC0892. KAKUSAN. DATAのデータをユーザーのデータセットへ写すことにより解決されますので御了承願います。

Members	Group	No. of entries	No. of bases
GENE1	Mammalian	889	704008
GENE2	Other Vertebrate	215	149972
GENE3	Invertebrate	263	179215
GENE4	Plant	154	151635
GENE5	Organelle	152	221561
GENE6	Bacteria	342	385089
GENE7	Structural RNA	286	29252
GENE8	Virus	587	599575
GENE9	Bacteriophage	80	169180
GENE10	Synthesis	43	13109
SUM		3012	2602596

表1 CC0892. KAKUSAN. DATAのメンバー名と収録されているデータ

2. 利用におけるコマンドの説明

このデータベースの利用者は計算機センターへ利用申請（手続は無料）したユーザーでなければならない。使用の時、まずLOGONコマンドにより、計算機を起動させるとREADYが表示される。以下のコマンドR1、R2、…等は、READYモードで働くコマンドである。

(R1) PFD┘2

と入力後図2の様に入力すると、画面は図3の様になりメンバー名が表示される。

今、メンバーGENE2について見る時はカーソルを移動させ、GENE2の左にSを入力すると、図4の様に CC0892. KAKUSAN. DATA (GENE2) が画面に表示される。

データの検索について、今、CCGGの配列を調べる時、F┘CCGGと入力しPF5を押せばCCGGを探してくれる。PFDを終了させる時はCAN (CANCEL) と入力すれば図3の画面にかわる。次にPF3を押し続けるとREADYにもどる。

メンバーGENE1、GENE6、GENE8については、データ量が多いためPFDコマンドは使用できない。メンバーGENE1、GENE6、GENE8については、(R2) のEDITコマンドを使用して下さい。

(R2) E┘'CC0892.KAKUSAN.DATA(GENE1)'

と入力すると、データセットタイプを要求してくるから

```

00000010 MEMBER                1589 BP    MRNA                PRE-ENTRY 12/05/83
00000020 CROSS REF              GENBANK: BLYAMYA
00000030 DEFINITION              BARLEY ALPHA-AMYLASE GENE MRNA.
00000040 ACCESSION                J01236
00000050 SOURCE                    BARLEY (HORDEUM VULGARE L. CV. HIMALAYA 1979 CROP; ALEURONE
00000060 CELL STIMULATED WITH GIBBERELIC ACID), CDNA TO MRNA, CLONE
00000070 E.
00000080 REFERENCE                    1 (BASES 1 TO 1589)
00000090 AUTHORS                      ROGERS, J. C. AND MILLIMAN, C.
00000100 TITLE                        ISOLATION AND SEQUENCE ANALYSIS OF A BARLEY ALPHA-AMYLASE
00000110 CDNA CLONE
00000120 JOURNAL                        J BIOL CHEM 258, 8169-8174 (1983)
00000130 COMMENT                        1 SUGGESTS THAT THERE ARE TWO ALPHA-AMYLASE GENES IN BARLEY
00000140 ALEURONE CELLS, AND THAT EXPRESSION OF THESE GENES IS
00000150 DIFFERENTLY EFFECTED BY GIBBERELIC ACID..
00000160 ORIGIN: 94 BP UPSTREAM OF NCOI SITE.
00000170 SEQUENCE                      345 A    484 C    480 G    280 T
00000180 ATCAGTCGATCAGCCAGTCAGCCAATCCTTCATTCATCCGGAGAAGAAAAGGAGTGTCTT      60
00000190 GTACTGTAAAGTGAGACAATCGACAGTAGCGCGCCATGGGGAAGAACGGCAGCCTGTGC      120
00000200 TGCTTCTCTCTGCTGCTGCTGCTGCTTCTCGCCGGTGGCGTCCGGCCACCAAGTCCTC      180
00000210 TTTCAAGGGTTCAACTGGGAGTCGTGGAAGCAGAGCGCGGGTGGTACAACATGATGATG      240
00000220 GGCAAGGTCGACGACATCGCCGCTGCGGAGTACCCACGCTGCGCTGGCCACCGCGTCG      300
00000230 CACTCCGTCTCCAACGAAGGTTACATGCCTGGTCCGGCTGTACGACATCGACGCGTCCAAG      360
00000240 TACGGCAACGCGCGGAGCTCAAGTCGCTCATCGGCGCGCTCCACGGCAAGGGCGGTGAG      420
00000250 GCCATCGCCGACATCGTCAATCAACCACCGCTGCGCGACTACAAGGATAGCCGCGGCATC      480
00000260 TACTGCATCTTCGAGGGCGGCACCTCCGACGGCCCGCTCGACTGGGGCCCCACATGATC      540
00000270 TGTCCGACGACACCAATACTCCGATGGCAGCCGCAAACTCGACACCGGAGCCGACTTC      600
00000280 GCGCGCGCGCCGACATCGACCACCTCAACGACCGGGTCCAGCGCGAGCTCAAGGATGG      660
00000290 CTCCTCTGGCTCAAGAGCGACCTCGGCTTCGACGCGTGGCGCCTTGACTTCGCTAGGGGC      720
00000300 TACTCGCCGGAGATGGCCAAGGTGTACATCGACGGCACATCCCCGAGCCTCGCCGTGGCC      780
00000310 GAGGTGTGGGACAATAATGGCCACCGCGCGGACGGCAAGCCCAACTACGACCGAGCAGCG      840
00000320 CACCGCAGAATCTGGTGAACCTGGTGGCAAGGTGGGCGGCGCGGCTCGGACGGCATG      900
00000330 GTGTTGACTTCAACGCAAAAGGATACTGAACGCTGCCGTGGAGGGCGAGCTGTGGAGG      960
00000340 CTGATCGACCCCGCAGGGGAAGGCCCGCGCTGATGGGATGGTGGCGGCAAGGCCGCC      1020
00000350 ACCTTCGTGCAACACCAGTACAGGCTCCACGCAGGCCATGTGGCCATTCCCTCCGAC      1080
00000360 AAGGTCATGCAGGGCTACGCGTACATCCTCACCCACCCGGCATCCCATGCATCTTCTAC      1140
00000370 GACCATTTCTTCAACTGGGGTTAAGGACCAGATCGCGGCGCTGGTGGCGATCAGGAAG      1200
00000380 CGCAACGGATCACGGCGACGAGCGCTCTGAAGTCTCATGCACGAAGGAGATGCCTAC      1260
00000390 GTCGCGGATAGCAGCGCAAGGTGGTGAAGTCCGGTCCAGGTACGACGCTCGGGGCG      1320
00000400 GTGATCCCGCGCGGTTCTGTGACCTCGGCACACGGCAACGACTACGCGTCTGGGAGAAG      1380
00000410 AACGGTGC CGCGGCAACACTACAACGGAGCTGAAGTCTGCACTGATCCGTCATTGATCG      1440
00000420 AGCATGAATTTCTGAAGTACATGATTCACCTTCTGGTATTACGCGGATATGATTAAC      1500
00000430 TGTATACCTGTACCCAAATTGATGAGCTATACGAATTGCTCACAGACTGTAGTAAATA      1560
00000440 AAGTACTGAGGATACGAAGAGGTTAATA:                                     1589
00000450 //
00000460 MEMBER                152 BP    MRNA                PRE-ENTRY 09/01/83
00000470 CROSS REF              GENBANK: BLYB1HOR
00000480 DEFINITION              BARLEY B1 HORDEIN MRNA (PARTIAL).
00000490 ACCESSION                J01237
00000500 KEYWORDS                HORDEIN; HOR-2 LOCUS; ENDOSPERM.
00000510 SOURCE                    BARLEY(HORDEUM VULGARE L. CV. SUNDANCE)CDNA TO ENDOSPERM
00000520 MRNA.
00000530 REFERENCE                    1 (BASES 1 TO 152)
00000540 AUTHORS                      FORDE, B. G., KREIS, M., BAHRAMIAN, M. B., MATTHEWS, J. A., MIFLIN,
00000550 B. J., THOMPSON, R. D., BARTELS, D. AND FLAVELL, R. B.
00000560 TITLE                        MOLECULAR CLONING AND ANALYSIS OF CDNA SEQUENCES DERIVED
00000570 FROM POLYA+ RNA FROM BARLEY ENDOSPERM: IDENTIFICATION OF B
00000580 HORDEIN RELATED CLONES
00000590 JOURNAL                    NUCL ACID RES 9, 6689-6707 (1981)
00000600 SEQUENCE                      42 A    32 C    35 G    43 T
00000610 CTCAACTTGAGGCGACAACCTTCCATTGCGTTGCGTACCCTACCAACGATGTGCGAGTGT      60
00000620 ATGTGCCGTTGTACCGCATAGTGCCATTAGCCATTGACACCAGAGTTGGTTLCTAATGAT      120
00000630 AAGAAAAGGGTCTTCTAGAAATATATAGTTGA:                               152
00000640 //

```

図1 CC0892. KAKUSAN. DATA (GENE4) の一部

```

-----< EDIT - DATASET MENU >-----
SPECIFY DATASET TO BE EDITED:

PFD LIBRARY DATASET:
PROJECT ==> CC0892
LIBRARY ==> KAKUSAN ==> ==> ==>
TYPE ==> DATA
MEMBER ==> (BLANK TO DISPLAY MEMBER LIST)

OTHER PARTITIONED OR SEQUENTIAL DATASET:
DATASET NAME ==>
VOLUME SERIAL ==> (IF NOT IN CATALOG)

```

図2 PFDモードにおけるEDIT

```

EDIT - MEMBER LIST - CC0892.KAKUSAN.DATA -----
COMMAND ==> SCROLL ==> PAGE
NAME VER.MOD CREATED LAST MODIFIED SIZE INIT MOD ID
GENE1
GENE10
GENE2
GENE3
GENE4
GENE5
GENE6
GENE7
GENE8
GENE9
**END**

```

図3 CC0892.KAKUSAN.DATAのメンバーリスト

DATA

と入力すると

E

と表示され、EDITモードになる。

FS

と入力すると、フルスクリーンとなり、画面に CC0892.KAKUSAN.DATA (GENE1) が表示される。図5にEDITコマンドの使用例を示す。

データ検索にはFINDコマンドを用いると便利で、今GCGCを探す時、F/GCGCとすればよい。

```

EDIT --- CC0892.KAKUSAN.DATA(GENE2) ----- COLUMNS 001 072
COMMAND ---> SCROLL ==> HALF
***** *****TOP OF DATA *****V10L10*****
--MSG> **WARNING**STATS MODE OF PROFILE IS SET TO ON. THIS MEMBER DOES NOT HAVE
--MSG> STATISTICS YET, BUT THEY WILL BE GENERATED WHEN DATA IS SAVED.
000003 MEMBER 482 BP DNA ENTRED 10/03/83
000004 CROSS REF GENBANK: CAIIGHV
000005 DEFINITION CAIMAN IG H-CHAIN V GENE HOMOLOGUE.
000006 ACCESSION J00804
000007 KEYWORDS IMMUNOGLOBULIN; HEAVY CHAIN; VARIABLE REGION.
000008 SOURCE CHIMAN CROCODYLUS GENOMIC DNA.
000009 REFERENCE 1 (BASES 1 TO 482)
000010 AUTHORS LITMAN,G.W., BETGER,L., MURPHY,K., LITMAN,R., HINDS,K.,
000011 JAHN,C.L. AND ERICKSON,B.W.
000012 TITLE COMPLETE NUCLEOTIDE SEQUENCE OF AN IMMUNOGLOBULIN VH GENE
000013 HOMOLOGUE FROM CAIMAN, A PHYLOGENETICALLY ANCIENT REPTILE
000014 JOURNAL NATURE 303, 349-352 (1983)
000015 COMMENT THE CAIMAN SEQUENCE HAS 70% HOMOMOLOGY WITH THE MURINE VH GENE
000016 PROBE (T15 FAMILY) USED TO ISOLATE IT. MOST OF THE HOMOMOLOGY
000017 OCCURS IN THE FRAMEWORK REGIONS (FR), WHILE THE HIGHLY
000018 VARIABLE REGIONS LACK EXTENSIVE HOMOMOLOGY. THE AUTHORS
000019 CONCLUDE THAT THE CLONE SEQUENCED DOES NOT REPRESENT A VH
000020 GENE FROM THE T15 FAMILY, BUT IS RATHER A VHIII HOMOLOGUE
000021 HAVING EXTENSIVE SIMILARITY IN THE FR GENE SEGMENTS. THE

```

図4 CC0892. KAKUSAN. DATA (GENE2) の一部

```

READY
E *CC0892.KAKUSAN.DATA(GENE1) *
KEQ52567A ENTER DATA SET TYPE-
DATA
E
FS
EDIT-FSO (V01/L06) --- CC0892.KAKUSAN.DATA(GENE1)
==>
ROW SCROLL ==> PAGE COLUMN SCROLL ==> 40 NONULLS 50
-----*-----1-----*-----2-----*-----3-----*-----4-----*-----5-----*-----6-----*-----7-
000010 MEMBER 611 BP ENTERED 06/10/82
000020 CROSS REF GENBANK: BOVAVPNPII
000030 DEFINITION BOVINE ARG. VASOPRESSIN & NEUROPHYSIN II PRECURSOR MRNA.
000040 ACCESSION J00001
000050 SOURCE BOVINE HYPOTHALAMUS. PLASMID PVNPII-1.
000060 REFERENCE 1 (BASES 1 TO 611)
000070 AUTHORS LAND,H., SCHUTZ,G., SCHMALE,H. AND RICHTER,D.
000080 TITLE NUCLEOTIDE SEQUENCE OF CLONED CDNA ENCODING BOVINE ARGININE
000090 VASOPRESSIN-NEUROPHYSIN II PRECURSOR
000100 JOURNAL NATURE 295, 299-303 (1982)
000110 COMMENT THE HYPOTHALAMIC NONAPRPTIDE HORMONE ARGININE VASOPRESSIN
000120 (AVP) AND ITS CARRIER PROTEIN NEUROPHYSIN II (NP11) ARE
000130 SYNTHESIZED FROM THE PRECURSOR GIVEN HERE. THE
000140 GLYCOPOLYPEPTIDE NOTED HAS BEEN ISOLATED FROM PITUITARIES OF
000150 SEVERAL SPECIES. BP 1-40 NOT FIRMLY ESTABLISHED.
000160 ORIGIN: 5' END OF MANA
000170 FEATURES FROM TO DESCRIPTION
000180 PEPT 50 550 ARG VASOPRESSIN-NEUROPHYSIN II PRECURSOR
000190 SIGP 53 106 SIGNAL PEPTIDE (PUTATIVE)
000200 MATP 107 133 ARGININE VASOPRESSIN

```

図5 EDITコマンドの使用方法

ここで、FINDコマンドは、行当たりについて検索するコマンドであり、2行にわたり調べるもの

でないということに注意していただきたい。EDITコマンド終了の時 END␣Nと入力するとREADYにもどる。

(R3) L␣'CC0892.KAKUSAN.DATA(GENE1)'

LISTコマンドによりデータを画面に表示させる時、途中でやめたい時はALTキーとDUPキーを同時に押すとREADYにもどる。

(R4) L␣'CC0892.KAKUSAN.DATA(GENE1)'␣100␣300

(R3)と同様。ただし行番号100から300までを表示する時。

(R5) LIST␣'CC0892.KAKUSAN.DATA(GENE2)'␣SY(C)

データCC0892.KAKUSAN.DATA(GENE2)を計算機センター1階のプリンターに出力する時。

(R6) L␣'CC0892.KAKUSAN.DATA(GENE2)'␣100␣300␣SY(C)

(R5)と同様。ただし行番号100から300までを出力。

(R7) COPY␣'CC0892.KAKUSAN.DATA(GENE3)'␣MY.DATA(GENE3)

KAKUSAN.DATA(GENE3)をデータセットMY.DATA(GENE3)にコピー。

(R8) COPY␣'CC0892.KAKUSAN.DATA'␣MY.DATA

(R7)と同様。ただし、データセット全体の複写。

(R9) M␣'CC0892.KAKUSAN.DATA(GENE5)'␣MY.DATA(GENE5)

(R7)と同様。

(E1) X␣L␣'CC0892.KAKUSAN.DATA(GENE6)'␣100␣300

EDITコマンドのFSモードで今、(R4)の機能を割込みによりおこなう時。

(E2) M␣'CC0892.KAKUSAN.DATA(GENE6)'␣100␣300␣500

上の(E1)によりデータを確認後MERGEコマンドにより KAKUSAN.DATA(GENE6)の100から300行を画面の500行の後に導入する時。

(R10) DEL␣MY.DATA

データセットMY.DATA全体を消す時。

(R11) DEL␣MY.DATA(GENE7)

データセット MY.DATAのメンバーGENE7を消す時。

(R12) LISTC

データセット名を表示する時。

(R13) LISTD␣MY.DATA␣M

データセット MY.DATAのメンバー名を表示

(R14) DS␣M/MY.DATA

(R13)と同様

(R15) DSPRINT_MY.DATA(GENE7)PI01

データセット MY. DATA (GENE7) をプリンターPI01に出力する時。

(R16) PR_P/MY.DATA(GENE7)

(R15)と同様。もよりのプリンターに出力する時。

(R17) PR_C

プリンター出力をキャンセルする時。

以上のコマンドの詳しい説明は計算機センター発行の利用の手引き MNJV01、MNHV01、MNIV01、MNKV02を見ていただきたい。

KAKUSAN. DATAの入出力に関するプログラムを図6に示す。

```
00000010 ***      MAIN PROGRAM
00000020 ***      PRINT THE NAME OF GENE DATA
00000030          PARAMETER (NMAX=50000 )
00000040          CHARACTER P1*4, P2*6, P3*62
00000050          CHARACTER SEQ(NMAX)
00000060          K1=0
00000070          LOC=0
00000080          NBP=0
00000090      101  CONTINUE
00000100          READ(5, '(A4,A6,A62)', END=100) P1, P2, P3
00000110          LOC=LOC+1
00000120          IF(P1.EQ.'MEMB') THEN
00000130              K1=K1+1
00000140              READ(P3, '(14X,15)') NUM
00000150              NBP=NBP+NUM
00000160          ENDIF
00000170          IF(P1.EQ.'DEPI') THEN
00000180              WRITE(6, '(1H,A62,18)') P3, LOC*10
00000190          ENDIF
00000200          IF(P1.EQ.'SEQU') THEN
00000210              READ(5, '(60A1)') (SEQ(K), K=1,NUM)
00000220              CALL SEQPRI(NUM,SEQ)
00000230          ENDIF
00000240          GO TO 101
00000250      100  CONTINUE
00000260          WRITE(6,*) 'NO. OF DATA AND SUM OF BP', K1, NBP
00000270          END
00000280          SUBROUTINE SEQPRI(NUM,SEQ)
00000290          ***      PRINT SEQUENCE OF GENE
00000300          PARAMETER (NMAX=50000 )
00000310          CHARACTER SEQ(NMAX)
00000320          WRITE(6,601)
00000330      601  FORMAT(1H,6(4X,'+',4X'*') )
00000340          LINE=(NUM-1)/60+1
00000350          DO 10 J=1,LINE
00000360              J1=60*(J-1)+1
00000370              J2=60*J
00000380              IF(NUM.LE.J2) J2=NUM
00000390              WRITE(6, '(1H,60A1)') (SEQ(K),K=J1,J2)
00000400              WRITE(6, '(1H+.60X,17)') J2
00000410      10  CONTINUE
00000420          RETURN
00000430          END
```

図6 CC0892. KAKUSAN. DATAのデータの入出力を扱うプログラム

このプログラムはFORTRAN 77で書いたものであり、CC0892. KAKUSAN. DATA (GENE 4) をデータとして実行した結果の一部を図7に示す。

```

BARLEY ALPHA-AMYLASE GENE MRNA.
+ * + * + * + * + * + *
ATCAGTCGATCAGCCAGTCAGCCAATCCTTCATTCATCCGGAGAAGAAAAGGAGTGTTC 60
GTACTGTAAGTGAGACAATCGACAGTAGCGCGCCATGGGGAAGAACGGCAGCCTGTGC 120
TGCTTCTCTCTGCTGCTGCTGCTGCTTCTCGCGGGTGGCGTCCGGCCACCAAGTCCCTC 180
TTTCAGGGGTCAACTGGGAGTCGTGGAAGCAGAGCGGCGGGTGGTACAACATGATGATG 240
GGCAAGGTCGACGACATCGCCGCTGCCGGAGTCACCCAGTCTGGCTGCCACCGCCGTCG 300
CACTCCGTCTCCAACGAAGTTACATGCCCTGGTGGCTGTACGACATCGACGCGTCCAAG 360
TACGGCAACGCGGCGGAGCTCAAGTCGCTCATCGGCGCGCTCCACGGCAAGGGCGTGCAG 420
GCCATCGCCGACATCGTCAATCAACCACCGCTGCGCCGACTACAAGGATAGCCGCGGCATC 480
TACTGCATCTTCGAGGGCGGCACCTCCGACGGCCCGCTGACTGGGGCCCCACATGATC 540
TGTCGCGACGACACCAAACTCTCGATGGCACCGCAAACCTCGACACCGGAGCCGACTTC 600
GCCGCGCGCCGACATCGACCACCTCAACGACCGGGTCCAGCGCGAGCTCAAGGAGTGG 660
CTCCTCTGGCTCAAGAGCGACCTCGGCTTCGACGCGTGGCGCTTGACTTCGCTAGGGGC 720
TACTCGCCGGAGATGGCCAAGGTGTACATCGACGGCACATCCCGAGCCTCGCGTGGCC 780
GAGGTGTGGGACAATATGGCCACCGGCGGCGACGGCAAGCCCAACTACGACAGGACGGCG 840
CACCGGCAGAATCTGGTGAATGGGTGGACAAGGTGGGCGGCGCGGCCTCGGCAGGCATG 900
GTGTTGACTTCAGGACCAAAAGGATACTGAACGCTGCCGTGGAGGGCGAGCTGTGGAGG 960
CTGATCGACCCCGAGGGGAAGGCCCGCGGTGATGGGATGGTGGCGGCCAAGGCCGCC 1020
ACCTTCGTCGACAACACGATACAGGCTCCACGCGAGGCCATGTGGCCATTCGCCCTCCGAC 1080
AAGGTCATGCAGGGCTACGCGTACATCCTCACCCACCCCGGCATCCCATGCATCTTCTAC 1140
GACCATTTCTTCAACTGGGGGTTAAAGGACAGATCGCGGCGCTGGTGGCGATCAGGAAG 1200
CGCAACGGCATCAGCGGACGAGCGCTCTGAAGATCCTCATGCACGAAGGAGATGCCATC 1260
GTCGCGGAGATAGACGGCAAGGTGGTGGTGAAGATCGGGTCCAGGTACGACGTCGGGGCG 1320
GTGATCCCGCGGGTTCGTGACCTCGGCACACGGCAACGACTACGCCGCTCGGGAGAAG 1380
AACGGTCCCGCGGCAACACTACAACGGAGCTGAAGTCTGCACTGATCCGTCAATCGATCG 1440
AGCATGAATTTCTGAAGTACATGATTCACTTCTGGTATTACGCGGATATGATTAACATA 1500
TGTATACCTGTACCCAAAATTGATGAGCTATACGAATTGCTCACAGCACTGTAGTAAATA 1560
AAGTACTGAGGATACGAAGAGTTAATA 1589
BARLEY B1 HORDEIN MRNA (PARTIAL).
+ * + * + * + * + * + *
CTCAACTTGAGGGCGACAACCTTCCATTGCGTTGCGTACCCTACCAACGATGTGCAGTGTTA 60
ATGTGCCGTTGTACCGCATAGTGCCATTAGCCATTGACACCAGAGTTGGTGTCTAATGAT 120
AAGAAAAGGGTCTTCTAGAAATATATAGTTGA 152

```

図7 CC0892. KAKUSAN. DATA (GENE 4) のデータを用い、図6のプログラムを実行した結果の一部

データのDEFINITIONが印刷され、行番号も出力される。続いて、塩基配列が出力される。塩基配列の印刷は行番号220のCALL SEQPRI (NUM, SEQ) によっておこなわれている。今、行番号220をコメント扱いにすれば、データのリストと行番号が得られる。CC0892. KAKUSAN. DATA (GENE 1) について実行した結果を図8に示す。

プログラムを準備してデータ処理をおこなうにはコントロール文が必要であり、この時利用の手引きMNIV01、MNKV01、MNQV01を見ていただきたい。また、データをフロッピーにいれ、マイコンで使用される場合、フロッピーの利用手引を参照していただきたい。

BOVINE ARG. VASOPRESSIN & NEUROPHYSIN II PRECURSOR MRNA.	30
BOVINE CHYMOSIN A (RENNIN) MRNA.	380
BOVINE CHYMOSIN B (RENNIN) MRNA.	710
BOVINE CHYMOSIN C (RENNIN) MRNA (PARTIAL).	1060
BOVINE ADRENAL PREPROENKEPHALIN MRNA.	1410
BOVINE FIBRINOGEN BETA-CHAIN MRNA.	2010
BOVINE ANTI-HAEMOPHILIC FACTOR IX MRNA.	2140
BOVINE GROWTH HORMONE (PRESOMATOTROPIN) GENE AND FLANKS.	2340
BOVINE PITUITARY GLYCOPROTEIN HORMONES PRE-ALPHA SUBUNIT	2990
BOVINE LOW MOLECULAR WEIGHT PREKININOGEN MRNA TYPE1.	3160
.	
.	
.	
.	
RAT ALPHA-TUBULIN MRNA (CLONE PT25), 3' END.	212800
RAT ALPHA-TUBULIN GENE, EXON 1 AND FLANKS.	213070
RAT ALPHA-TUBULIN GENE, EXONS 2,3 & 4 AND 3' FLANK.	213460
RAT ALPHA-TUBULIN PSEUDOGENE AND FLANKS.	213930
RAT (SPRAGUE-DAWLEY) U1 RNA LOCUS A GENE.	214270
RAT U2 SMALL NUCLEAR RNA GENE AND FLANKS.	214440
RAT U2 SMALL NUCLEAR RNA PSEUDOGENE.	214560
RAT WHEY ACIDIC PROTEIN MRNA.	214680
RAT WHEY PHOSPHOPROTEIN MRNA CLONE.	215000
SHEEP CORTICOTROPIN-RELEASING FACTOR (CRF) PRECURSOR MRNA.	215280
NO. OF DATA AND SUM OF BP	889 704008

図8 CC0892. KAKUSAN. DATA (GENE 1) のデータのリストを出力させた結果

おわりに

データベース登録に関して教えていただいた経済学部 山下邦弘先生に感謝します。

参考マニュアル

- MNHV01 関崎 正夫著
TSS端末によるバッチ処理・—フルスクリーン機能と出力検索—
- MNIV01 山崎 光悦著
TSSコマンド、データセット、実行のABC
- MNJV01 中島 恵美著
FORTRANユーザのためのTSS入門
- MNKV01 車古 正樹著
TSSユーザのための便利なコマンド(その1)
- MNKV02 車古 正樹著
TSSユーザのための便利なコマンド(その2)
- MNQV01 山下 邦弘著
FORTRANユーザのためのジョブ制御文とデータセット

文 献

1. 江口 至洋、陶山 明、上田 裕三、和田 昭允：蛋白質核酸酵素、28、923-942 (1983)。
2. 陶山 明、江口 至洋、上田 裕三、和田 昭允：蛋白質核酸酵素、28、1165-1186 (1983)。
3. 久原 哲、榊 佳之、高木 利久、松尾 文碩、二村 祥一、鬼塚 千代子：九州大学大型計算機センター広報、16、497-521 (1983)。
4. 小澤 宏：東京大学大型計算機センターニュース、15、59-68 (1983)。
5. 金久 実：細胞工学、2、1520-1529 (1983)。