ＵＷＧＣＧ及びＩＤＥＡＳ 利用の手引

1988年 6月17日

国立遺伝学研究所ＤＤＢＪ利用者講習会資料

藤田 信之

411 静岡県三島市
国立遺伝学研究所
分子遺伝研究部門

電話： 0559-75-0771
E-mail: nfujita@niguts.nig.junet

Sequence Comparison Programs

BESTFIT	makes optimal alignments by method of Smith/Waterman
GAP	makes optimal alignments by method of Needleman/Wunsch
COMPARE-DOTPLOT	makes a dot plot by method of Maizel/Lenk
WORDSEARCH -SEGMENT	makes rapid sequence comparisons (for data base searching)
CONSENSUS	creates a consensus table from pre-aligned sequences
FITCONSENSUS	finds sequences similar to a know consensus sequence

Mapping

MAP	displays restriction sites and possible translations
MAPSORT	tabulates maps sorted by fragment position and size
MAPPLOT	show restriction maps graphically
FIND	finds short sequence patterns in sequences
PEPTIDEMAP	show patterns of proteolytic cleavage
PEPTIDESORT	tabulates data on peptides from proteolytic cleavage

Nucleic Acids Secondary Structure

STEMLOOP	finds possible stem (inverted repeat) and loop structure
COMPARE-DOTPLOT	makes a dot plot by method of Maizel/Lenk

Protein Secondary Structure

PEPPLOT	plots measures of protein secondary structure

Pattern Recognition and Composition Analysis

TESTCODE	finds possible coding regions by plotting the "TestCode" statistic of Fickett
CODONPREFERENCE	plots the similarity between the codon choices in each reading frame and a codon frequency table
FRAMES	plots positions of rare codons and open reading frames
STATPLOT	plots the frequencies of any pattern in a DNA sequence
COMPOSITION	measures composition and di and trinucleotide frequencies
FINGERPRINT	shows labelled fragments expected for an RNA fingerprint
REPEAT	finds direct repeats
CODONFREQUENCY	tabulates codon frequencies
CORRESPOND	finds similar patterns of codon choice by comparing codon frequency tables
TERMINATOR	uses method of Trifonov to find prokaryotic terminators

Sequence Manipulation

SEQED	screen-oriented sequence editor for entering, editing and checking sequences
ASSEMBLE	joins sequences together
REFORMAT	converts a sequence file from one format to another
FROMEMBL	converts a file from EMBL format to UWGCG format

FROMGENBANK	converts a file from GenBank format to UWGCG format
FROMSTADEN	converts a file from STADEN format to UWGCG format
TOSTADEN	converts a file from UWGCG format to STADEN format

Sequence Conversion

BACKTRANSLATE	translates a peptide into a nucleotide sequence
TRANSLATE	translates a nucleotide into a peptide sequence
REVERSE	reverses and/or complements a sequence
SHUFFLE	randomizes a sequence while maintaining its composition
SIMPLIFY	makes a simplification scheme for peptide sequence comparisons

Printing and Publication

PUBLISH	arranges sequences for publication
POSTER	arranges text on figures for publication or for poster sessions
OVERPRINT	prints sequence figures using a daisy wheel printer

Database Search

STRINGS	finds files by character string (key word)
NAMES	finds files by names
SHOWFILES	writes a documented "file of file names"
FETCH	copies UWGCG data file(s) into your directory

I D E A S

Database Management and Simple Analysis Program

S E Q M A N Sequence database manipulation program

g e t	get entries from library
s c a n	scan contents (identifiers and definition) of library
f i n d	find entries by matching pattern search in definition
l i s t	list entries by searching identifiers
x r e f	identify corresponding entries in two databases
p r	print sequence
r c	print reverse complement
t r	translate to amino acids
d i f f	difference of two sequences
f r e q	count residue or codon frequency
o r f	locate open reading frames
a n n	annotate sequence
s r c h	search sequence pattern in sequence or library
c v s	convert to the library (UWGCG) format
s g c	define special genetic code
t t y	set output line width
h e l p	read help files
j o i n	join lines of two files
u c	convert to upper-case letters
l c	convert to lower-case letters
t y p e	type and format lines
d i r	directory (similar to the DCL command)
d c l	temporarily return to the DCL level
e n d	escape from SEQMAN

Sequence Homology Search Programs

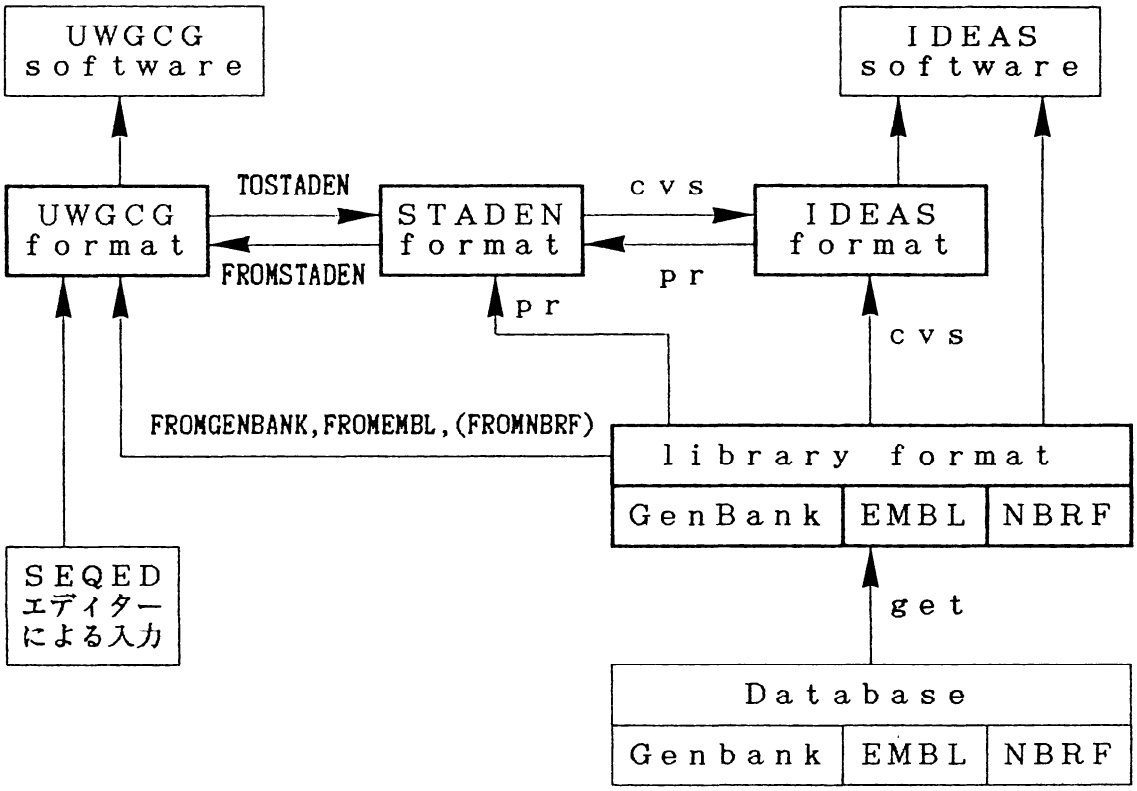
S E Q H	local homology in nucleic acid sequences (complete version)
S E Q F	local homology in nucleic acid sequences (rapid version)
S E Q L	locally stable socondary structure in nucleic acid
S E Q A	global homology alignment of two sequences
S E Q H P	local homology in protein sequences (complete version)
S E Q D P	significance of homology in protein sequences
S E Q F P	local homology in protein sequences (rapid version)
S E Q F T	local homology in translated protein sequences

Protein Structure/Function Prediction Programs

S T R A L I	secondary structure prediction by homology alignment
C H O F A S	secondary strucuture prediction by Chou-Fasman's method
D E L P H I	Garnier's program for secondary structure prediction
D S S P	Kabsch-Sander's program for secondary structure assigmment
A N N O T	annotation of secondary structure in the Protein Data Bank
H P L O T	distribution of hydrophobic and charged amino acids
H C O M P	comparison of hydrophobicity profiles in two proteins
W U K A B	Wu-Kabat plot of amino acid variability
A L O M	prediction of membrane proteins and transmembrane segments

フォーマット変換

個々のデータベースおよび解析ソフトウェアがそれぞれ異なったデータフォーマットを採用しているため、フォーマット変換（データの加工）をする必要がかなり頻繁に生じます。 下に変換のおおまかな流れ図を示しますが、ＩＤＥＡＳやＵＷＧＣＧのプログラムを使いこなすには、この過程を十分に理解しておくことが重要です。



上の図の中で、"ｇｅｔ", "ｃｖｓ", "ｐｒ"はＩＤＥＡＳ／ＳＥＱＭＡＮの中のコマンドを表わし、"TOSTADEN","FROMSTADEN","FROMGENBANK","FROMEMBL", "FROMNBRF"はＵＷＧＣＧの中のプログラムを表わす。

getコマンド（ＩＤＥＡＳ／ＳＥＱＭＡＮ）の使い方

```
          genbank
  %get    embl        sequenceID >outputfile
          nbrf
          genbank
  %get    embl        <listfile >outputfile
          nbrf
```

cvsコマンド（ＩＤＥＡＳ／ＳＥＱＭＡＮ）の使い方

```
  %cvs  inputfile  >outputfile
```

prコマンド（ＩＤＥＡＳ／ＳＥＱＭＡＮ）の使い方

```
  %pr  inputfile  >outputfile
```

例題１：ＧｅｎＢａｎｋ ＤＮＡデータベースより必要な配列を取り出し、ＵＷＧＣＧ上の"ＭＡＰ"を用いて制限酵素の切断点を検索する。

```
nigvms$ naq genbank[CR]  . . . . . 1
Nucleic Acid Query System V3/2.4

NAQ> find e. coli heat shock[CR]  . . . . . 2
ECOC625     E.coli heat-inducible promoter for operon encoding heat shock
              protein C62.5
ECODNAK     E.coli dnaK gene encoding the heat shock 70 protein
ECOHTPR     E. coli htpR (heat shock regulatory) gene, complete cds
ECOHTPRR    E.coli htpR gene coding for heat shock regulatory protein F33.4 and
              livJ gene promoter
M12565      E. coli K12 dnaJ+ gene encoding a heat shock protein
  5 entries found

NAQ> quit[CR]  . . . . . 3

nigvms$ ideas seqman[CR]  . . . . . 4

% get genbank ecohtpr >ecohtpr.[CR]  . . . . . 5

% type ecohtpr.[CR]  . . . . . 6
LOCUS       ECOHTPR      1060 bp ds-DNA             entered   07/01/85
DEFINITION  E. coli htpR (heat shock regulatory) gene, complete cds.
ACCESSION   K02177
KEYWORDS    heat shock protein.
                .
                .
                .

% end[CR]  . . . . . 7

nigvms$ uwgcg[CR]  . . . . . 8

                        Welcome to the
            University of Wisconsin Genetics Computer Group
                        Version 3.0, June 1985
              Help is available with the command $ GenHelp

nigvms$ fromgenbank[CR]  . . . . . 9

 Reformat what GenBank data file?  ecohtpr.[CR]  . . . . . 10
       ecohtpr.  1060 bp.
 reformatted: ecohtpr.
 total files: 1
 total bases: 1060

nigvms$ type ecohtpr.[CR]  . . . . . 11
locus       ecohtpr      1060 bp ds-dna             entered   07/01/85
definition  e. coli htpr (heat shock regulatory) gene, complete cds.
accession   k02177
keywords    heat shock protein.
                .
                .
                .
```

1．データベースを検索するためＮＡＱを呼び出す．検索は
ＧｅｎＢａｎｋ ＤＮＡデータベースを対象とする．

2．Ｄｅｆｉｎｉｔｉｏｎ行に"Ｅ．"，"ｃｏｌｉ"，"ｈｅａｔ"，
"ｓｈｏｃｋ"を含むエントリーを検索する．

3．ＮＡＱから抜け出る．

4．必要なデータを取り出すためＩＤＥＡＳ中のＳＥＱＭＡＮを
呼び出す．　単に"ＩＤＥＡＳ↵"とのみ入力するとメニュー
が表示されるので，その中からＳＥＱＭＡＮを指定するこ
も可能．

5．ＧＥＴコマンドにより，ＧｅｎＢａｎｋデータベース中のエ
ントリー"ＥＣＯＨＴＰＲ"を取り出し，自分のディレクトリ
に格納する．ファイル名は"ＥＣＯＨＴＰＲ．"とする．

6．確かにコピーできたかどうかをＴＹＰＥコマンドで確認する．

7．ＳＥＱＭＡＮから抜け出る．同時にＩＤＥＡＳからも抜け出
る．

8．ＵＷＧＣＧの使用を宣言する．（最初に一度のみ必要）

9．ＧｅｎＢａｎｋフォーマットからＵＷＧＣＧフォーマットに
変換するためＦＲＯＭＧＥＮＢＡＮＫを呼び出す．

10．フォーマット変換するファイルの名前を入力する．

11．確かにフォーマット変換できたかどうかをＴＹＰＥコマンド
で確認する．

```
nigvms$ map[CR] . . . . . 12

MAP is a sequence display tool that shows the sequence and its complement
with the restriction map above and possible protein translations below.
Local data files:  Enzyme.dat (enzyme names and recognition sites)
Command Line Switches
        /SIX            shows six-base cutters when enzyme name is "*"
        /ONCe           shows only enzymes that cut once in the chosen range
        /LINear         treats the sequence as linear (default)
        /CIRcular       treats the sequence as circular
        /WIDth          allows choice of number of characters per line
        /MISmatch       allows mismatches in site search
        /APPend         appends the enzyme data file to MAP's output

 (Linear) MAP of what sequence ?  ecohtpr.[CR]  . . . . . 13
                Begin (* 1 *) ?  1[CR]  . . . . . 14
              End (*  1060 *) ?  120[CR]  . . . . . 15
 Enter enzymes:  Use a "*" to get all, a "#" to get none
 or enter the names individually, one per line, ending
 the list with a blank line.
                  Enzyme(* * *):  *[CR]  . . . . . 16
 What protein translations do you want:
      a) frame 1   b) frame 2   c) frame 3
      d) frame 4   e) frame 5   f) frame 6
      t)hree forward frames   s)ix frames   o)pen frames only
      n)o protein translation   q)uit
 Please select (capitalize for 3-letter) (* t *):  to[CR]  . . . . . 17
 What should I call the output file (* ecohtpr.map *) ?  term[CR]  . . . . . 18

 MAP of: ecohtpr.  check: 6042  from: 1  to: 120
                          13-JUN-88 11:33  ..

         H
         I A
         N L
         D U
         3 1

         AAGCTTGCATTGAACTTGTGGATAAAATCACGGTCTGATAAAACAGTGAATGATAACCTC
       1 ---------+---------+---------+---------+---------+---------+ 60
         TTCGAACGTAACTTGAACACCTATTTTAGTGCCAGACTATTTTGTCACTTACTATTGGAG

a        K  L  A  L  N  L  W  I  K  S  R  S  D  K  T  V  N  D  N  L
b         S  L  H  *                                    M  I  T  S
c           A  C  I  E  L  V  D  K  I  T  V  *

                                                H          E
                 M    A                        I HH       C CMT  M
                 N    L                        N HA       O LNA  B
                 L    U                        P AE       R ALQ  O
                 1    1                        1 12       V 111  2
                                                            //
         GTTGCTCTTAAGCTCTGGCACAGTTGTTGCTACCACTGAAGCGCCAGAAGATATCGATTG
      61 ---------+---------+---------+---------+---------+---------+ 120
         CAACGAGAATTCGAGACCGTGTCAACAACGATGGTGACTTCGCGGTCTTCTATAGCTAAC

a        V  A  L  K  L  W  H  S  C  C  Y  H  *
b         L  L  L  S  S  G  I  V  V  A  T  T  E  A  P  E  D  I  D  *
c

 nigvms$
```

１２．制限酵素の切断点を求めるためＭＡＰを呼び出す．

１３．ファイル名を入力する．

１４．データの始まりを指定する．

１５．データの終わりを指定する．

１６．制限酵素名を入力する．”＊”を指定すると登録されている酵素すべてが対象となる．

１７．アミノ酸配列への翻訳を３つのフレームについて行なう．ただしオープンフレームのみを表示する．

１８．結果を端末（ＴＥＲＭ）に出力する．ファイル名を指定してファイルに書き出すことも可能．

例題２：画面エディターを用いてアミノ酸配列を入力し，ＩＤＥＡＳ上の"ＳＥＱＦＰ"により
ＮＢＲＦ－ＰＩＲデータベースを対象としてホモロジー検索を行なう．

```
nigvms$ uwgcg[CR]  . . . . . 1
                    Welcome to the
          University of Wisconsin Genetics Computer Group
                    Version 3.0, June 1985
           Help is available with the command $ GenHelp

nigvms$ seqed[CR]  . . . . . 2

What sequence to edit ?  test.seq[CR]  . . . . . 3
```

```
SeqEd                              test.seq                              UWGCG
         :test sequence[CR]  . . . . . 4
         :[CR]
         :


    MADKMQSLALAPVGGLDSIYRAANAWPMLSADEERALAEKL  . . . . . 5
....l.........l.........l.........l.........l.........l.........l.........l....
    0         10        20        30        40        50        60        70

    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    l......l......l......l......l......l......l......l......l......l......l
    0      10     20     30     40     50     60     70     80     90     100
```

```
:check[CR]  . . . . . 6
```

```
SeqEd                              test.seq                              UWGCG
         :test sequence
         :
         :


    MADKMQSLALAPVGGLDSYIRAANAWPMLSADEERALAEL  . . . . . 7
                                             ^
    MADKMQSLALAPVGGLDSYIRAANAWPMLSADEERALAEKLHYHGDLEAAKTLILSHLRVVHIARNYAGYGLPQ
....l.........l.........l.........l.........l.........l.........l.........l....
    0         10        20        30        40        50        60        70

    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    l......l......l......l......l......l......l......l......l......l......l
    0      50     100    150    200    250    300    350    400    450    500
```

```
:exit[CR]  . . . . . 8

nigvms$ type test.seq[CR]  . . . . . 9

test sequence
        test.seq  Length: 284  13-JUN-88 19:48  Check: 7288  ..
      1  MADKMQSLAL APVGGLDSYI RAANAWPMLS ADEERALAEK LHYHGDLEAA
     51  KTLILSHLRF VVHIARNYAG YGLPQADLIQ EGNIGLMKAV RRFNPEVGVR
    101  LVSFAVHWIK AEIHEYVLRN WRIVKVATTK AQRKLFFNLR KTKQRLGWFN
    151  QDEVEMVARE LGVTSKDVRE MESRMAAQDM TFDLSSDDDS DSQPMAPVLY
    201  LQDKSSNFAD GIEDDNWEEQ AANRLTDAMQ GLDERSQDII RARWLDEDNK
    251  STLQELADRY GVSAERVRQL EKNAMKKLRA AIEA
```

1．ＵＷＧＣＧの使用を宣言する（最初に一度だけ必要）．

2．配列編集用の画面エディター"ＳＥＱＥＤ"を呼び出す．
（ＳＥＱＥＤを用いるためには，ＶＴエミュレーターが必要）

3．作成するファイルの名前を入力する．

4．コメントを入力する．コメントは５行まで入力できるが，他のフォーマットに変換する際失われることに注意．

5．配列を大文字で入力する．画面の範囲を越えると自動的にスクロールされる．

主な編集コマンド	
→，←	一文字右へ，一文字左へ
ｎ→，ｎ←	ｎ文字右へ，ｎ文字左へ
＞，＜	５０文字右へ，５０文字左へ
：ｎ[CR]	ｎ番目の文字へ
／文字列[CR]	文字列の検索
[ＤＥＬ]	一文字削除
:n1，n2ｄｅｌｅｔｅ[CR]	n1番目からn2番目まで削除

6．コロン ":"に続けて "ｃｈｅｃｋ[CR]"と入力し，再入力モードに入る．

7．入力の誤りをさがすため，配列を再度入力する．一致しない箇所には "^"が表示されブザーが鳴るので修正する．行の間の移動は "↓"および "↑"で行なう．

8．確認が終わったらコロン ":"に続けて "ｅｘｉｔ[CR]"と入力し，エディターから抜け出す．

9．ファイルが出来ていることをＴＹＰＥコマンドで確認する．

```
nigvms$ tostaden[CR]  . . . . . 10

TOSTADEN writes a UWGCG sequence file into a file in the Staden/Sanger
format. If the file contains a nucleic acid sequence, the ambiguity codes
are translated as shown in Appendix III.

 TOSTADEN of what UWGCG sequence ? test.seq[CR]  . . . . . 11
                Begin (* 1 *) ?  1[CR]  . . . . . 12
              End (*  284 *) ?  284[CR]  . . . . . 13
 What should I call the output file (* test.sdn *) ?  test.seq[CR]  . . . . . 14

nigvms$ type test.seq[CR]  . . . . . 15

MADKMQSLALAPVGGLDSYIRAANAWPMLSADEERALAEKLHYHGDLEAA
KTLILSHLRFVVHIARNYAGYGLPQADLIQEGNIGLMKAVRRFNPEVGVR
LVSFAVHWIKAEIHEYVLRNWRIVKVATTKAQRKLFFNLRKTKQRLGWFN
QDEVEMVARELGVTSKDVREMESRMAAQDMTFDLSSDDDSDSQPMAPVLY
LQDKSSNFADGIEDDNWEEQAANRLTDAMQGLDERSQDIIRARWLDEDNK
STLQELADRYGVSAERVRQLEKNAMKKLRAAIEA

nigvms$ ideas[CR]  . . . . . 16

                        I D E A S
          Integrated Database and Extended Analysis System
                  for Nucleic Acids and Proteins
                      Dr. Minoru Kanehisa
                Institute of Chemical Research
                        Kyoto University
                      0774-32-3111 ext. 2222

          Enter "Menu" to see the description of programs.
          Enter "Help" to access on-line documentation.

  Options available:
   SEQMAN  FRAMIS  SEQF    SEQFN   SEQH    SEQFP   SEQHP   SELECT  GRAPH
   SEQDP   SEQFT   SEQA    SEQL    SEQG    STRALI  CHOFAS  DELPHI  DSSP
   ANNOT   HPLOT   HCOMP   WUKAB   ALIGN   ALOM
   Symbol  Manual  Help    Menu    Exit
Option? seqman[CR]  . . . . . 17

% cvs test.seq >test.seq[CR]  . . . . . 18

% type test.seq[CR]  . . . . . 19

;    284 residues
TEST.SEQ
MADKMQSLAL APVGGLDSYI RAANAWPMLS ADEERALAEK LHYHGDLEAA KTLILSHLRF
VVHIARNYAG YGLPQADLIQ EGNIGLMKAV RRFNPEVGVR LVSFAVHWIK AEIHEYVLRN
WRIVKVATTK AQRKLFFNLR KTKQRLGWFN QDEVEMVARE LGVTSKDVRE MESRMAAQDM
TFDLSSDDDS DSQPMAPVLY LQDKSSNFAD GIEDDNWEEQ AANRLTDAMQ GLDERSQDII
RARWLDEDNK STLQELADRY GVSAERVRQL EKNAMKKLRA AIEA1

% end[CR]  . . . . . 20

  Options available:
   SEQMAN  FRAMIS  SEQF    SEQFN   SEQH    SEQFP   SEQHP   SELECT  GRAPH
   SEQDP   SEQFT   SEQA    SEQL    SEQG    STRALI  CHOFAS  DELPHI  DSSP
   ANNOT   HPLOT   HCOMP   WUKAB   ALIGN   ALOM
   Symbol  Manual  Help    Menu    Exit
Option? seqfp[CR]  . . . . . 21

IDEAS Sequence Analysis System -- Protein Database Search
```

１０．UWGCGフォーマットからＳＴＡＤＥＮフォーマット（中間フォーマット）に変換するためＴＯＳＴＡＤＥＮを呼び出す．

１１．変換するファイルの名前を入力する．

１２．データの始まりを指定する．

１３．データの終わりを指定する．

１４．変換後のデータを格納するファイルの名前を入力する．

１５．フォーマット変換できていることをＴＹＰＥコマンドで確認する．

１６．ＩＤＥＡＳを呼び出す．

１７．ＩＤＥＡＳの中のＳＥＱＭＡＮを呼び出す．

１８．ＣＶＳコマンドにより中間フォーマットの”ＴＥＳＴ．ＳＥＱ”をＩＤＥＡＳフォーマットに変換する．変換後のデータは”ＴＥＳＴ．ＳＥＱ”の名前でファイルに書き出す．

１９．フォーマット変換できていることをＴＹＰＥコマンドで確認する．

２０．ＳＥＱＭＡＮから抜け出る．

２１．ホモロジー検索を行なうためＩＤＥＡＳの中のＳＥＱＦＰを呼び出す．

```
Input file: test.seq[CR]  . . . . . 22
Library file ([CR] for NBRF database) [CR]  . . . . . 23
Output file ([CR] for your terminal) test.dat[CR]  . . . . . 24
Change parameters? (Y/N/H) y[CR]  . . . . . 25
Enter MAXD,KTPL,NW,LWID,DEL,BIAS (Use / to keep default values)
-70/[CR]  . . . . . 26
Parameters are:
  MAXD = -70    KTPL = 2    NW = 10    LWID =  80
  DEL =  8    BIAS =  0
Sequence or END: test.seq[CR]  . . . . . 27
Start,End ([CR] for entire sequence) [CR]  . . . . . 28
Do you want to submit a batch job? (Y/N) y[CR]  . . . . . 29

*** File SEQFP.COM created ***
FORTRAN STOP
Batch queue? [SYS$BATCH]
Job SEQFP (queue SYS$BATCH, entry 702) started on SYS$BATCH

  Options available:
   SEQMAN  FRAMIS  SEQF    SEQFN   SEQH    SEQFP   SEQHP   SELECT  GRAPH
   SEQDP   SEQFT   SEQA    SEQL    SEQG    STRALI  CHOFAS  DELPHI  DSSP
   ANNOT   HPLOT   HCOMP   WUKAB   ALIGN   ALOM
   Symbol  Manual  Help    Menu    Exit
Option? exit[CR]  . . . . . 30

nigvms$ type test.dat[CR]  . . . . . 31

QUERY SEQUENCE  TEST.SEQ    FROM      1  TO    284  TOTAL   284

>RNECS       613 amino acids                      0      1     13     18 H
DNA-directed RNA polymerase (EC 2.7.7.6) sigma chain - Escherichia coli  T+
       10         20         30         40         50         60         70
ADKMQSLALAPVGGLDSYIRAANAWPMLSADEERALAEKLHY-HGDLE-AAKTLILSHLRFVVHIARNYAGYGLPQADLI
  :          :          :                         : :      :: :  ::  :   ::   :::
SEKLHDVSEEVHRALQKLQQIEEE-TGLTIEQVKDINRRMSIGEAKARRAKKEMVEANLRLVISIAKKYTNRGLQFLDLI
 330        340        350        360        370        380        390        400
         90        100        110        120        130        140        150
QEGNIGLMKAVRRFNPEVGVRLVSFAVHWIKAEIHEYVLRNWRIVKVATTKAQ--RKLFFNLRKTKQRLGWFNQDEVEMV
::::::::::: :    :      : ::   :        :             ::    :   : :       :
QEGNIGLMKAVDKFEYRRGYKFSTYATWWIRQAITRSIADQARTIRIPVHMIETINKLNRISRQMLQEMGR-EPTPEELA
  410        420        430        440        450        460        470        480
160        170        180        190        200        210        220        230
ARELGVTSKDVREMESRMAAQDMTFDLSSDDDSDSQPMAPVLYLQDKSSNFADGIEDDNWEEQAANRLTDAMQGLDERSQ
 : :       :     :           :: ::         :                   ::    :   :: :
ERML-MPEDKIRKVL-KIAKEPISMETPIGDDEDSH-LGD--FIEDTTLELPLDSATTE-SLRAATH--DVLAGLTAREA
    490        500        510        520        530        540        550
240        250        260        270
DIIRARW-LDEDNKSTLQELADRYGVSAERVRQLEKNAMKKLR
   : :   :     :: :      :  :: :: : :  : :::
KVLRMRFGIDMNTDYTLEEVGKQFDVTRERIRQIEAKALRKLR
 560        570        580        590

      HOMOLOGY SCORE =  -312    ( -161 )
      PERCENT MATCH  =  25.1%   (  71 / 283 )
             .
             .
             .

THE BEST HOMOLOGY SCORE (UNCORRECTED)  -1355  WAS FOUND IN:
  RGECH

nigvms$
```

２２．検索するファイルの名前を入力する．

２３．検索の対象となる配列集合を指定する．

２４．検索結果の出力先を指定する．

２５．検索のパラメーターを変更する．

２６．最初のパラメーター（ＭＡＸＤ）を－７０に変更する．２番目
以降のパラメーターは初期設定のままとする．

２７．データのＩＤを入力する．

２８．データの始まりと終わりを指定する．

２９．以上の設定をバッチジョブとして実行させる．

３０．ＩＤＥＡＳから抜け出る．

この段階で一旦ＬＯＧＯＵＴするか，別の作業を行なう．

３１．バッチジョブが終わったら結果を出力させる．