<http://www.bioinformatics.org/wiki/Hidden_Markov_Model>

Hidden Markov Model

Markov chains are named for Russian mathematician [Andrei Markov](http://en.wikipedia.org/wiki/Andrei_Markov) (1856-1922), and they are defined as observed sequences. A Markov model is a system that produces a Markov chain, and a hidden Markov model is one where the rules for producing the chain are unknown or "hidden." The rules include two probabilities: (i) that there will be a certain observation and (ii) that there will be a certain state transition, given the state of the model at a certain time.[[1]](http://www.bioinformatics.org/wiki/Hidden_Markov_Model#cite_note-dugad-0)

The Hidden Markov Model (HMM) method is a mathematical approach to solving certain types of problems: (i) given the model, find the probability of the observations; (ii) given the model and the observations, find the most likely state transition trajectory; and (iii) maximize either *i* or *ii* by adjusting the model's parameters. For each of these problems, algorithms have been developed: (i) [Forward-Backward](http://www.bioinformatics.org/w/index.php?title=Forward-Backward&action=edit&redlink=1), (ii) [Viterbi](http://www.bioinformatics.org/w/index.php?title=Viterbi&action=edit&redlink=1" \o "Viterbi (page does not exist)), and (iii) [Baum-Welch](http://www.bioinformatics.org/w/index.php?title=Baum-Welch&action=edit&redlink=1) (and the [Segmental K-means](http://www.bioinformatics.org/w/index.php?title=Segmental_K-means&action=edit&redlink=1) alternative).[[1]](http://www.bioinformatics.org/wiki/Hidden_Markov_Model#cite_note-dugad-0)[[2]](http://www.bioinformatics.org/wiki/Hidden_Markov_Model#cite_note-1)

The HMM method has been traditionally used in signal processing, speech recognition, and, more recently, [bioinformatics](http://www.bioinformatics.org/wiki/Bioinformatics). It may generally be used in [pattern recognition](http://www.bioinformatics.org/w/index.php?title=Pattern_recognition&action=edit&redlink=1) problems, anywhere there may be a model producing a [sequence of observations](http://www.bioinformatics.org/w/index.php?title=Sequence_of_observations&action=edit&redlink=1). In bioinformatics, it has been used in [sequence alignment](http://www.bioinformatics.org/wiki/Sequence_alignment), [*in silico* gene detection](http://www.bioinformatics.org/w/index.php?title=In_silico_gene_detection&action=edit&redlink=1), [structure prediction](http://www.bioinformatics.org/wiki/Structure_prediction), [data-mining literature](http://www.bioinformatics.org/w/index.php?title=Data-mining_literature&action=edit&redlink=1), and so on.

Here is a simple example of the use of the HMM method in *in silico* gene detection:

* [Codons](http://www.bioinformatics.org/w/index.php?title=Codons&action=edit&redlink=1) (or [DNA triplets](http://www.bioinformatics.org/w/index.php?title=DNA_triplets&action=edit&redlink=1)) are the observations.
  + 64 codons equals a probability of 1/64.
* The DNA sequence is the Markov chain (set of observations).
* Switches from one genomic region to another are the state transitions.
  + 5'[UTR](http://www.bioinformatics.org/w/index.php?title=UTR&action=edit&redlink=1) to [CDS](http://www.bioinformatics.org/w/index.php?title=CDS&action=edit&redlink=1)
  + CDS to [intron](http://www.bioinformatics.org/w/index.php?title=Intron&action=edit&redlink=1)
  + intron to CDS
  + CDS to 3'UTR
  + 4 state transitions equals a probability of ¼.

Difficulties with the HMM method include the need for accurate, applicable, and sufficiently sized [training sets](http://www.bioinformatics.org/w/index.php?title=Training_set&action=edit&redlink=1) of data. As for the example of gene detection, in order to accurately predict genes in the [human genome](http://www.bioinformatics.org/w/index.php?title=Human_genome&action=edit&redlink=1), many genes in the genome must be accurately known.

<http://www.genome.jp/tools/motif/hmmformat.htm>

## Profile data format

### 1. Pfam / HMMER Profile file format

Here we describe briefly about data formats used in **Pfam** database and **HMMER** application. Precise documentation should be found in original [HMMER site](http://hmmer.wustl.edu/).

The profile HMM calculated from multiple sequnce alignment data in this service is stored in Profile HMM **save format** (usually with ".hmm" extension). It is an ASCII file containing a lot of header and descriptive records followed by large numerical matrix which holds probabilistic model of the motif. The file of this format is useful to search against sequnce databases to find out other proteins which share the same motif.   
This HMM file should not be edited manually (especially the matrix part) because it contains consistent numerical model as a whole.

On the other hand another data can be shown after searching Pfam database using a protein sequence as a query or by keyword search against Pfam in DBGET. This data format is called **SELEX** and is a kind of multiple sequence alignment format. It is also an ASCII file with defined header information (lines start with "#=") followed by aligned sequences with their accesion numbers.   
This file can be edited manually to append and remove the sequences or to extract interesting part of multiple sequence alignment. Additionally this file can be converted to HMM matrix file using our service then be used to retrieve the sequence databases.

### 2. Sample of the Pfam/HMMER profile data formats

#### Profile HMM save file format

HMMER2.0

NAME aligment15900

LENG 30

ALPH Amino

RF no

CS no

MAP yes

COM /bio/package/motif2/tool/hmmer/hmmer-2.1.1-sgi-irix/binaries/hmmbuild /var/tmp/hmmbuild15900.out /var/tmp/aligment15900

COM /bio/package/motif2/tool/hmmer/hmmer-2.1.1-sgi-irix/binaries/hmmcalibrate --cpu 16 --seed 0 /var/tmp/hmmbuild15900.out

NSEQ 7

DATE Sat Sep 11 15:05:30 1999

CKSUM 5380

XT -8455 -4 -1000 -1000 -8455 -4 -8455 -4

NULT -4 -8455

NULE 595 -1558 85 338 -294 453 -1158 197 249 902 -1085 -142 -21 -313 45 531 201 384 -1998 -644

EVD -18.948267 0.332753

HMM A C D E F G H I K L M N P Q R S T V W Y

m->m m->i m->d i->m i->i d->m d->d b->m m->e

-263 \* -2585

1 771 -1882 -411 2059 -2183 -1618 -445 -1810 -132 -1938 -1108 1307 -1797 -45 -623 -660 -691 833 -2234 -1612 1

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 -263 \*

2 -3359 -3855 4091 -1760 -4660 -2944 -2744 -5137 -3226 -4926 -4562 -2110 -3493 -2619 -3804 -3177 -3522 -4662 -4114 -4101 2

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

3 -1065 -2468 1603 -57 -2947 -1668 -722 -2722 -525 -2693 -1839 2792 -1991 -320 -1100 512 1054 -2243 -2892 -2154 3

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

4 -727 -1948 -540 -89 -2262 -1651 2840 -1971 74 -1998 -1146 -366 959 2104 -346 -652 1144 -1602 -2216 -1605 4

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

5 -538 -1629 -562 1017 -1818 -1578 -258 -1476 1251 -1588 1411 -289 -1668 139 -379 848 602 610 -1902 -1314 5

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

6 1314 -906 -1359 855 1195 -1842 -641 -470 -657 -801 -110 -901 -1934 -528 -959 -845 1146 696 -1338 -912 6

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

7 -1971 -3909 2965 1196 -4068 -1863 -1228 -4005 -1500 -3867 -3211 2561 -2406 -900 -2382 -1619 -2042 -3473 -4062 -3022 7

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

8 -2369 -1902 -4793 -4363 -1231 -4499 -3691 3398 -4119 1125 -63 -4176 -4130 -3594 -3972 -3819 -2322 373 -2826 -2644 8

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

9 -869 -1300 -2621 -2496 -2154 -1852 -2151 1399 -2305 -1899 -1440 -1983 -2448 -2165 -2378 3040 -1167 -933 -2713 -2253 9

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

10 1329 -1161 -1275 -735 -1431 -1731 -706 1139 -512 -1261 -534 -845 -1925 -474 1165 1430 -620 -769 -1729 -1277 10

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

11 -4031 -3046 -4530 -4815 2343 -4401 2189 -2979 -4380 -2358 -2406 -3049 -4266 -3170 -3789 -3636 -3891 -3109 4356 3468 11

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

12 -3904 -3244 -4483 -4744 4436 -3934 -1848 -2962 -4575 -2394 -2507 -3766 -4186 -3845 -4126 -4011 -3988 -3213 -1146 -58 12

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

13 -613 -2085 1242 1129 -2395 -1530 -232 -2146 159 -2097 1115 841 858 1854 -356 -487 -557 -1702 -2271 -1582 13

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

14 -701 -2129 1299 2032 -2423 -1580 -304 -2160 86 -2129 1903 -219 -1721 137 1105 -577 -646 -1739 -2319 -1642 14

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

15 2466 -1028 -2804 -2415 -1627 -1982 -1861 1566 -2171 -1267 -744 -1960 -2421 -1952 -2231 -1217 1287 -190 -2196 -1843 15

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

16 735 -1007 -1545 -997 -1140 -1930 -808 2134 1030 -935 -275 -1060 -2060 -681 -991 -971 1197 -293 -1536 -1118 16

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

17 976 -2990 2423 1857 -3293 460 -850 -3108 -773 -3037 -2220 -369 -2125 -465 -1424 -1153 -1385 -2629 -3228 -2403 17

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

18 -3966 -3031 -4498 -4772 4179 -4363 -705 -2844 -4348 -2217 -2283 -3069 -4249 -3180 -3779 -3629 -3846 -3016 41 2228 18

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

19 -2399 -1929 -4821 -4381 -1188 -4524 -3684 3341 -4137 1299 -16 -4200 -4133 -3576 -3972 -3840 -2347 317 -2793 -2638 19

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

20 -1716 -3017 3517 -361 -3889 -1860 -1457 -3862 -1732 -3830 -3148 -699 -2444 -1159 -2504 1207 -1902 -3217 -3941 -3075 20

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

21 272 1602 -560 -9 -1850 457 -257 -1515 1266 -1616 -782 -286 -1666 1451 -372 322 -479 -1207 2061 -1330 21

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

22 781 2726 -2853 -2284 -742 -2181 -1286 969 -1958 -554 52 -1864 -2328 -1646 -1877 733 -714 1920 -1272 -921 22

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

23 -820 -2085 1236 -136 -2390 -1744 -340 -2067 1873 947 -1197 -405 -1838 91 975 -724 -742 -1704 -2234 -1655 23

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

24 691 -2471 1487 987 -2765 -1629 -484 -2539 -198 -2476 -1589 1136 -1855 2499 -756 -765 -893 -2082 -2652 -1914 24

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -1720 -7442 -534 -894 -1115 -701 -1378 \* \*

25 -575 -1287 -74 22 -1375 -1107 -153 -1361 382 -1331 -825 -161 -1435 3145 180 -573 -610 -1149 -1476 -958 25

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -40 -5762 -6804 -894 -1115 -180 -3089 \* \*

26 260 1379 -479 76 -2234 -1554 -213 -1950 1621 -1939 -1046 1289 -1650 232 1267 490 -513 -1547 -2139 -1497 26

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

27 -1193 -2205 -753 -669 -3139 2350 -1057 -2890 -539 -2873 -2071 2508 -2238 -689 835 -1165 -1296 -2371 -2984 -2445 27

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

28 -2083 -3069 -616 1596 -4116 3207 -1886 -4090 -2021 -4049 -3402 -1234 -2789 -1620 -2577 -1970 -2275 -3484 -3912 -3428 28

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

29 -1994 -2859 -2322 -1319 -3595 -2641 -719 -3017 2195 -2743 -2025 -1327 1433 -295 2873 -1881 -1760 -2738 -2649 -2446 29

- -149 -500 233 43 -381 399 106 -626 210 -466 -720 275 394 45 96 359 117 -369 -294 -249

- -12 -7442 -8484 -894 -1115 -701 -1378 \* \*

30 -2174 -1982 -4186 -4170 -2484 -3406 -3713 42 -3993 -1584 -1517 -3697 -3771 -3891 -3924 -3057 -2329 3646 -3507 -3148 30

- \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \* \*

- \* \* \* \* \* \* \* \* 0

//

#### SELEX Multiple sequence alignment file format

#=GF ID Y\_phosphatase

#=GF AC PF00102

#=GF DE Protein-tyrosine phosphatase

#=GF AU Sonnhammer ELL

#=GF AL Clustalw

#=GF SE Swissprot\_feature\_table

#=GF GA -125 -125

#=GF TC -123.40 -123.40

#=GF NC -126.80 -126.80

#=GF BM hmmbuild -F HMM SEED

#=GF BM hmmcalibrate --seed 0 HMM

#=GF DR PROSITE; PDOC00323;

#=GF DR PRINTS; PR00700;

#=GF DR SCOP; 1ypt; fa;

#=GF DC The following Pfam-B families contain sequences that according to Prodom

#=GF DC are members of this Pfam-A family.

#=GF DR PFAMB; PB035686;

#=GF DR PFAMB; PB036047;

#=GF DR PFAMB; PB036157;

#=GF DR PFAMB; PB036312;

#=GF DR PFAMB; PB036563;

#=GF CC The following Prosite members are absent due to lack of

#=GF CC homology outside the active site: Swiss:Q05918, Swiss:P34442,

#=GF CC Swiss:P40479, Swiss:P24656, Swiss:P43078, Swiss:P38590.

#=GF SQ 38

CD45\_RAT/628-862 NQNKNRYVDILPYDYNRVELSEINGDA..GSTYINAS......YIDGFKEPRK....YIAAQGPRD....ETVDDFWKMIWE...QKATVIVMVTRCEEGNRNKCAEYWPCMEE.........GTRTFRDVVVTINDHKR...CPDYIIQK..LSIAHK...........KEKATGREVTHIQFTSWPDH..GVPEDPHLLLKLRRRVNAFSN...FFS..........GPIVVHCSAGVGRTGTYIGIDAMLESLEAEG..........KVDVYGYVVNLRRQRC.LMVQVEAQYILIHQALVE

#=GS CD45\_RAT/628-862 AC P04157

LAR\_DROME/1497-1728 NKSKNRYANVTAYDHSRVQLPAVEGVV..GSDYINAN......YCDGYRKHNA....YVATQGPLQ....ETFVDFWRMCWE...LKTATIVMMTRLEERTRIKCDQYWPTRG...........TETYGQIFVTITETQE...LATYSIRT..FQLCRQGFND............RREIKQLQFTAWPDH..GVPDHPAPFLQFLRRCRALTP...PES..........GPVIVHCSAGVGRTGCYIVIDSMLERMKHEK..........IIDIYGHVTCLRAQRN.YMVQTEDQYIFIHDAILE

#=GS LAR\_DROME/1497-1728 AC P16621

PTN1\_MOUSE/40-276 NKNRNRYRDVSPFDHSRIKLHQEDN......DYINAS......LIKMEEAQR....SYILTQGPLP....NTCGHFWEMVWE...QKSRGVVMLNRIMEKGSLKCAQYWPQQEEKE.......MVFDDTGLKLTLISEDV...KSYYTVRQ..LELENLTTK............ETREILHFHYTTWPDF..GVPESPASFLNFLFKVRESGSLSLEHG...........PIVVHCSAGIGRSGTFCLADTCLLLMDKRKDPS.......SVDIKKVLLEMRRFRM.GLIQTADQLRFSYLAVIE

#=GS PTN1\_MOUSE/40-276 AC P35821

----- (continue) -------

//

### 3. PROSITE Profile format

Aditionally here is the sample of PROSITE profile data format.

ID PARP\_ZN\_FINGER\_2; MATRIX.

AC PS50064;

DT NOV-1997 (CREATED); NOV-1997 (DATA UPDATE); JUL-1998 (INFO UPDATE).

DE Poly(ADP-ribose) polymerase zinc finger domain profile.

MA /GENERAL\_SPEC: ALPHABET='ABCDEFGHIKLMNPQRSTVWYZ'; LENGTH=89;

MA /DISJOINT: DEFINITION=PROTECT; N1=6; N2=84;

MA /NORMALIZATION: MODE=1; FUNCTION=LINEAR; R1=.1357; R2=.01506700; TEXT='NScore';

MA /CUT\_OFF: LEVEL=0; SCORE=555; N\_SCORE=8.5; MODE=1;

MA /CUT\_OFF: LEVEL=-1; SCORE=422; N\_SCORE=6.5; MODE=1;

MA /DEFAULT: D=-20; I=-20; B1=-40; E1=-40; MI=-105; MD=-105; IM=-105; DM=-105;

MA /I: B1=0; BI=-105; BD=-105;

MA /M: SY='F'; M=-20,-27,-23,-34,-27,64,-30,-7,0,-24,7,0,-20,-30,-30,-17,-20,-10,-3,16,46,-27;

MA /M: SY='A'; M=3,-13,-4,-16,-12,-20,3,-16,-22,-3,-13,-10,-8,-21,-10,2,-6,-10,-13,-24,-18,-12;

MA /M: SY='V'; M=20,-21,-13,-27,-21,-9,-19,-26,16,-17,3,3,-20,-20,-20,-22,-3,-2,25,-24,-13,-21;

MA /M: SY='E'; M=-11,14,-30,24,48,-32,-12,-1,-32,7,-22,-21,3,-3,16,-2,0,-11,-30,-30,-20,32;

MA /M: SY='Y'; M=-20,-20,-30,-20,-20,30,-30,20,0,-10,0,0,-20,-30,-10,-10,-20,-10,-10,30,80,-20;

MA /M: SY='A'; M=46,-9,-10,-18,-9,-20,0,-19,-11,-10,-12,-11,-8,-10,-9,-19,13,2,-1,-22,-20,-9;

MA /M: SY='K'; M=-11,-1,-30,-1,9,-29,-20,-9,-30,48,-29,-10,0,-11,10,34,-10,-10,-20,-20,-10,9;

MA /M: SY='S'; M=5,-1,-13,-2,-1,-19,-5,-10,-20,-5,-27,-18,8,-11,0,0,32,19,-10,-36,-18,-1;

MA /M: SY='G'; M=-1,5,-24,0,-9,-26,38,-10,-31,-12,-30,-20,17,-18,-10,-13,10,-8,-26,-29,-25,-10;

MA /M: SY='R'; M=-17,-9,-27,-10,-1,-19,-20,-3,-27,25,-19,-10,0,-19,7,60,-6,-2,-17,-21,-10,-1;

MA /M: SY='A'; M=37,-7,-10,-13,-7,-20,0,-17,-13,-10,-17,-13,-3,-10,-7,-17,20,7,-3,-27,-20,-7;

MA /M: SY='S'; M=7,-3,-13,-7,-7,-17,1,-15,-17,-11,-20,-13,3,-12,-6,-12,24,23,-8,-32,-17,-7;

MA /M: SY='C'; M=-10,-20,120,-30,-30,-20,-30,-30,-30,-30,-20,-20,-20,-40,-30,-30,-10,-10,-10,-50,-30,-30;

MA /M: SY='K'; M=-12,-2,-30,-2,8,-28,-20,-8,-30,47,-28,-10,0,-12,10,36,-10,-10,-20,-20,-10,8;

MA /M: SY='G'; M=-3,-5,-29,-5,-7,-29,30,-15,-35,8,-30,-16,1,-15,-7,0,-1,-14,-25,-21,-22,-7;

MA /M: SY='C'; M=-10,-20,120,-30,-30,-20,-30,-30,-30,-30,-20,-20,-20,-40,-30,-30,-10,-10,-10,-50,-30,-30;

MA /M: SY='K'; M=-7,-2,-27,0,14,-24,-7,-8,-23,16,-19,-6,-2,-11,6,6,-3,-8,-19,-24,-15,10;

MA /M: SY='E'; M=-9,12,-27,17,34,-31,-15,2,-26,6,-23,-16,8,-6,25,1,5,-6,-28,-30,-17,29;

MA /M: SY='K'; M=-5,3,-23,1,5,-25,-15,-9,-25,28,-28,-13,7,-11,5,16,4,3,-17,-26,-13,5;

MA /M: SY='I'; M=-10,-30,-30,-40,-30,0,-40,-30,50,-30,20,20,-20,-20,-20,-30,-20,-10,30,-20,0,-30;

MA /M: SY='E'; M=0,-1,-23,2,12,-21,-15,-11,-14,-6,-14,-13,-3,2,-2,-11,2,-3,-11,-31,-19,4;

MA /M: SY='K'; M=-4,-1,-28,-2,8,-29,-18,-11,-28,44,-28,-10,-1,-10,8,25,-8,-9,-18,-20,-11,8;

MA /M: SY='D'; M=-10,18,-30,28,4,-34,27,-9,-39,-9,-29,-24,9,-14,-8,-14,0,-15,-30,-30,-25,-2;

MA /M: SY='Q'; M=-2,-9,-19,-11,-3,-18,-19,-10,-3,-8,-8,-2,-5,-15,11,-8,7,6,-2,-28,-11,4;

MA /M: SY='L'; M=-8,-28,-3,-32,-25,2,-32,-25,22,-27,26,15,-26,-29,-22,-23,-21,-8,19,-26,-6,-25;

MA /M: SY='R'; M=-18,-10,-30,-10,-2,-21,-13,-2,-31,26,-21,-11,0,-20,8,63,-9,-11,-21,-20,-12,-2;

MA /M: SY='I'; M=-10,-29,-25,-35,-25,4,-34,-22,35,-27,31,26,-24,-24,-17,-24,-24,-10,20,-20,0,-24;

MA /M: SY='A'; M=23,-7,-16,-12,-7,-21,6,-14,-19,-5,-19,-14,-2,-13,-6,-4,13,1,-9,-25,-20,-7;

MA /M: SY='K'; M=-4,-12,-25,-15,-6,-14,-24,-18,-4,15,-10,-1,-10,-16,-5,5,-11,-8,0,-20,-7,-6;

MA /M: SY='M'; M=-7,-13,-20,-19,-12,-9,-21,-9,4,4,2,25,-13,-17,-3,-1,-10,1,6,-23,-5,-8;

MA /M: SY='V'; M=-1,-27,-11,-28,-27,-1,-27,-25,26,-18,9,15,-26,-28,-24,-18,-8,0,41,-29,-9,-26;

MA /M: SY='Q'; M=-9,-13,-26,-12,0,-14,-24,-6,-5,-5,-10,-1,-13,-5,14,-6,-7,-7,-6,-18,-3,6;

MA /M: SY='D'; M=-5,18,-19,22,5,-23,-6,3,-25,-6,-28,-21,17,-13,0,-8,18,5,-19,-35,-12,2;

MA /M: SY='P'; M=4,-15,-28,-13,-5,-23,-16,-21,-14,-11,-21,-15,-15,47,-11,-18,-1,2,-15,-28,-24,-11;

MA /M: SY='F'; M=-14,-9,-24,-10,7,14,-23,-7,-7,-9,0,7,-11,-16,-6,-10,-12,-10,-10,-15,2,2;

MA /M: SY='F'; M=-11,-13,-22,-16,-11,8,-22,7,-11,-3,-11,-3,-7,-21,-11,0,-6,-7,-5,-18,5,-11;

MA /I: I=-7; MI=-5; MD=-27; IM=-5;

MA /M: SY='P'; M=-5,-7,-17,-5,4,-11,-11,-5,-6,-3,-7,2,-8,19,0,-7,-6,-5,-10,-14,-10,1; D=-5;

MA /I: DM=-27;

MA /M: SY='Q'; M=-3,0,-13,0,7,-17,-8,0,-12,9,-13,-4,1,-5,17,6,3,-2,-12,-13,-7,12; D=-5;

MA /I: DM=-27;

MA /M: SY='D'; M=-15,12,-27,21,-1,-15,-8,-4,-19,-12,-5,-12,-1,-19,-9,-14,-11,-11,-17,-22,-1,-6;

MA /M: SY='G'; M=5,-10,-28,-11,-19,-29,62,-20,-37,-19,-28,-19,-1,-19,-19,-20,1,-18,-27,-20,-29,-19;

MA /M: SY='K'; M=-10,0,-27,0,-1,-24,-5,-8,-20,18,-17,3,-2,-14,2,7,-10,-12,-15,-23,-12,0;

MA /M: SY='I'; M=-6,-21,-21,-25,-16,-10,-29,-16,21,-14,6,16,-19,-21,-1,-14,-11,-6,20,-24,-6,-10;

MA /M: SY='P'; M=-13,0,-36,11,6,-33,-17,-14,-26,0,-30,-21,-7,52,-5,-11,-7,-10,-29,-31,-25,-2;

MA /M: SY='H'; M=-14,4,-16,0,2,-17,-17,24,-20,-4,-12,-7,12,-20,2,5,-7,-10,-21,-31,-6,0;

MA /M: SY='W'; M=-20,-40,-50,-40,-30,10,-20,-30,-20,-20,-20,-20,-40,-30,-20,-20,-40,-30,-30,150,30,-20;

MA /M: SY='Y'; M=-20,-16,-28,-18,-16,26,-27,34,-8,-14,-4,0,-12,-27,-11,-9,-17,-13,-14,9,53,-16;

MA /M: SY='H'; M=-20,0,-30,0,0,-20,-20,100,-30,-10,-20,0,10,-20,10,0,-10,-20,-30,-30,20,0;

MA /M: SY='P'; M=-11,-19,-29,-18,-7,-1,-27,-14,0,-9,-2,-2,-18,6,-12,-13,-15,-9,-5,-15,2,-11;

MA /M: SY='E'; M=-4,6,-22,8,13,-26,-10,-7,-25,7,-25,-16,5,-9,8,1,12,7,-18,-30,-16,11;

MA /M: SY='C'; M=-10,-20,120,-30,-30,-20,-30,-30,-30,-30,-20,-20,-20,-40,-30,-30,-10,-10,-10,-50,-30,-30;

MA /M: SY='F'; M=-19,-29,-20,-39,-29,70,-29,-17,3,-27,11,8,-20,-29,-35,-19,-20,-10,1,6,26,-27;

MA /M: SY='W'; M=-4,-30,-24,-34,-27,26,-23,-25,1,-22,-1,-4,-26,-27,-28,-20,-18,-11,5,33,12,-24;

MA /M: SY='K'; M=-8,1,-28,3,18,-32,-18,-2,-26,26,-26,-9,1,-9,26,16,0,-7,-23,-23,-12,22;

MA /I: MD=-27;

MA /M: SY='K'; M=-5,-4,-13,-4,-1,-8,-10,-6,-7,8,-3,-2,-2,-9,0,7,-5,-3,-6,-13,-5,-1; D=-5;

MA /I: I=-7; MI=0; IM=0; DM=-27;

MA /M: SY='R'; M=-15,-6,-26,-9,-3,-18,-19,-4,-20,22,-18,-8,2,-20,3,46,-8,-8,-11,-24,-11,-3;

MA /M: SY='A'; M=11,-1,-22,-2,7,-29,7,-9,-24,-4,-20,-13,-1,-11,8,-9,6,-7,-19,-24,-20,8;

MA /M: SY='R'; M=-16,-2,-30,2,23,-20,-20,6,-28,16,-19,-13,0,-12,13,31,-6,-11,-24,-23,-10,16;

MA /M: SY='L'; M=1,-21,-21,-21,-13,-7,-20,-20,5,-20,10,2,-19,0,-15,-19,-9,-4,2,-26,-12,-15;

MA /M: SY='G'; M=-2,-9,-22,-11,-15,-16,10,-20,-11,-18,-10,-8,-4,-17,-15,-18,3,3,-6,-26,-16,-16;

MA /M: SY='W'; M=-8,-17,-22,-20,-15,12,-16,-17,-15,-13,-15,-13,-10,-20,-14,-7,2,3,-11,14,5,-13;

MA /M: SY='H'; M=-7,-13,-23,-15,-7,-7,-22,8,-9,-1,-2,1,-9,-21,-2,6,-12,-10,-6,-20,0,-6;

MA /M: SY='P'; M=-3,-8,-31,-4,-4,-29,8,-17,-27,-5,-28,-18,-7,26,-10,-13,-3,-9,-25,-26,-25,-9;

MA /I: MD=-23;

MA /M: SY='E'; M=-2,4,-15,6,7,-15,-11,-8,-8,-5,-10,-9,1,-7,-1,-8,5,2,-6,-21,-10,2; D=-4;

MA /I: MD=-23;

MA /M: SY='Y'; M=-6,-10,-17,-10,-5,2,-3,-5,-6,-8,-5,-4,-9,-15,-9,-9,-6,-7,-3,-5,8,-8; D=-4;

MA /I: I=-6; MI=0; MD=-23; IM=0; DM=-23;

MA /M: SY='T'; M=0,2,-8,1,7,-10,-8,-6,-10,-2,-10,-8,1,-4,1,-4,10,13,-6,-17,-8,4; D=-4;

MA /I: DM=-23;

MA /M: SY='D'; M=-15,26,-29,35,27,-35,-14,9,-31,4,-25,-19,13,-9,18,-2,0,-10,-30,-33,-15,22;

MA /M: SY='L'; M=-9,-29,-22,-33,-25,4,-32,-23,31,-27,32,22,-26,-26,-20,-22,-23,-9,22,-21,-1,-24;

MA /M: SY='E'; M=-12,2,-30,8,19,-24,-20,-3,-21,6,-17,-9,-5,5,10,-2,-7,-10,-23,-23,-9,13;

MA /M: SY='G'; M=-2,0,-28,-4,-16,-28,56,-14,-36,-16,-30,-20,12,-20,-16,-16,2,-16,-30,-24,-28,-16;

MA /M: SY='F'; M=-19,-30,-26,-38,-29,59,-30,-19,3,-27,6,0,-23,-29,-33,-20,-23,-13,-1,26,31,-28;

MA /M: SY='S'; M=-1,0,-21,2,13,-25,-4,-8,-23,-4,-23,-16,2,-5,10,-6,16,7,-19,-31,-19,11;

MA /M: SY='E'; M=-7,5,-23,3,16,-17,-18,-7,-9,-4,-10,-10,7,-12,1,-8,2,0,-12,-31,-15,8;

MA /M: SY='L'; M=-10,-30,-21,-31,-21,9,-31,-21,23,-30,47,20,-29,-29,-20,-21,-29,-10,12,-20,0,-21;

MA /M: SY='R'; M=-5,-5,-24,-5,7,-22,-16,-6,-25,18,-20,-12,0,-13,7,32,1,-2,-17,-24,-13,4;

MA /M: SY='W'; M=-2,-8,-27,-9,-12,5,-14,-18,-18,-15,-14,-16,-12,-19,-16,-17,-9,-9,-15,25,4,-12;

MA /M: SY='E'; M=-7,23,-27,31,33,-31,-12,-1,-30,4,-23,-22,12,-7,8,-5,2,-8,-27,-33,-20,21;

MA /M: SY='D'; M=-19,45,-30,64,25,-39,-11,0,-39,1,-29,-29,17,-9,3,-9,0,-10,-30,-39,-20,14;

MA /M: SY='Q'; M=-10,0,-30,0,16,-36,-20,2,-24,25,-24,-4,0,-10,41,18,-4,-10,-26,-20,-10,28;

MA /M: SY='E'; M=-1,7,-27,12,36,-30,-16,-3,-26,11,-20,-15,0,-5,19,1,0,-8,-24,-27,-18,27;

MA /M: SY='K'; M=-4,1,-23,1,9,-24,-16,-6,-20,12,-18,-6,0,-11,10,9,1,2,-15,-25,-12,9;

MA /M: SY='I'; M=-8,-30,-21,-34,-26,4,-34,-26,34,-28,28,18,-26,-26,-22,-24,-21,-8,28,-22,-2,-26;

MA /M: SY='K'; M=-8,0,-25,-2,5,-25,-20,-12,-25,36,-25,-10,0,-10,5,21,-3,4,-15,-22,-10,5;

MA /M: SY='K'; M=-11,9,-29,9,12,-32,-17,-5,-29,35,-29,-12,7,-11,14,21,-6,-9,-23,-24,-12,13;

MA /M: SY='A'; M=8,-8,-20,-12,-3,-19,-15,5,-11,-4,-6,1,-6,-14,8,-6,-1,-2,-10,-22,-7,2;

MA /M: SY='I'; M=1,-27,-21,-32,-23,0,-29,-25,27,-26,25,14,-23,-23,-20,-24,-18,-7,21,-21,-4,-23;

MA /M: SY='P'; M=-2,-6,-32,1,23,-29,-17,-12,-23,-2,-23,-19,-10,39,2,-12,-3,-9,-26,-29,-24,11;

MA /M: SY='A'; M=12,4,-18,0,-6,-21,-6,-14,-14,-9,-16,-14,2,-4,-9,-14,7,6,-8,-29,-19,-8;

MA /M: SY='I'; M=-5,-23,-25,-27,-25,-8,-4,-25,14,-25,7,6,-16,-22,-21,-24,-12,-9,10,-22,-10,-25;

MA /M: SY='K'; M=-2,-3,-26,-3,1,-29,6,-11,-29,17,-28,-13,1,-12,6,7,2,-8,-21,-23,-17,4;

MA /I: E1=0; IE=-105; DE=-105;

NR /RELEASE=38,80000;

NR /TOTAL=18(10); /POSITIVE=18(10); /UNKNOWN=0(0); /FALSE\_POS=0(0);

NR /FALSE\_NEG=0; /PARTIAL=1;

CC /TAXO-RANGE=??E??; /MAX-REPEAT=2;

DR P49916, DNL3\_HUMAN, T; P97386, DNL3\_MOUSE, T; P18493, PPOL\_BOVIN, T;

DR P26446, PPOL\_CHICK, T; P35875, PPOL\_DROME, T; P09874, PPOL\_HUMAN, T;

DR P11103, PPOL\_MOUSE, T; P27008, PPOL\_RAT , T; Q11208, PPOL\_SARPE, T;

DR P31669, PPOL\_XENLA, T;

DR Q08824, PPOL\_ONCMA, P;

3D 2PAW; 1PAX; 2PAX; 3PAX; 4PAX; 1A26;

DO PDOC00360;

//