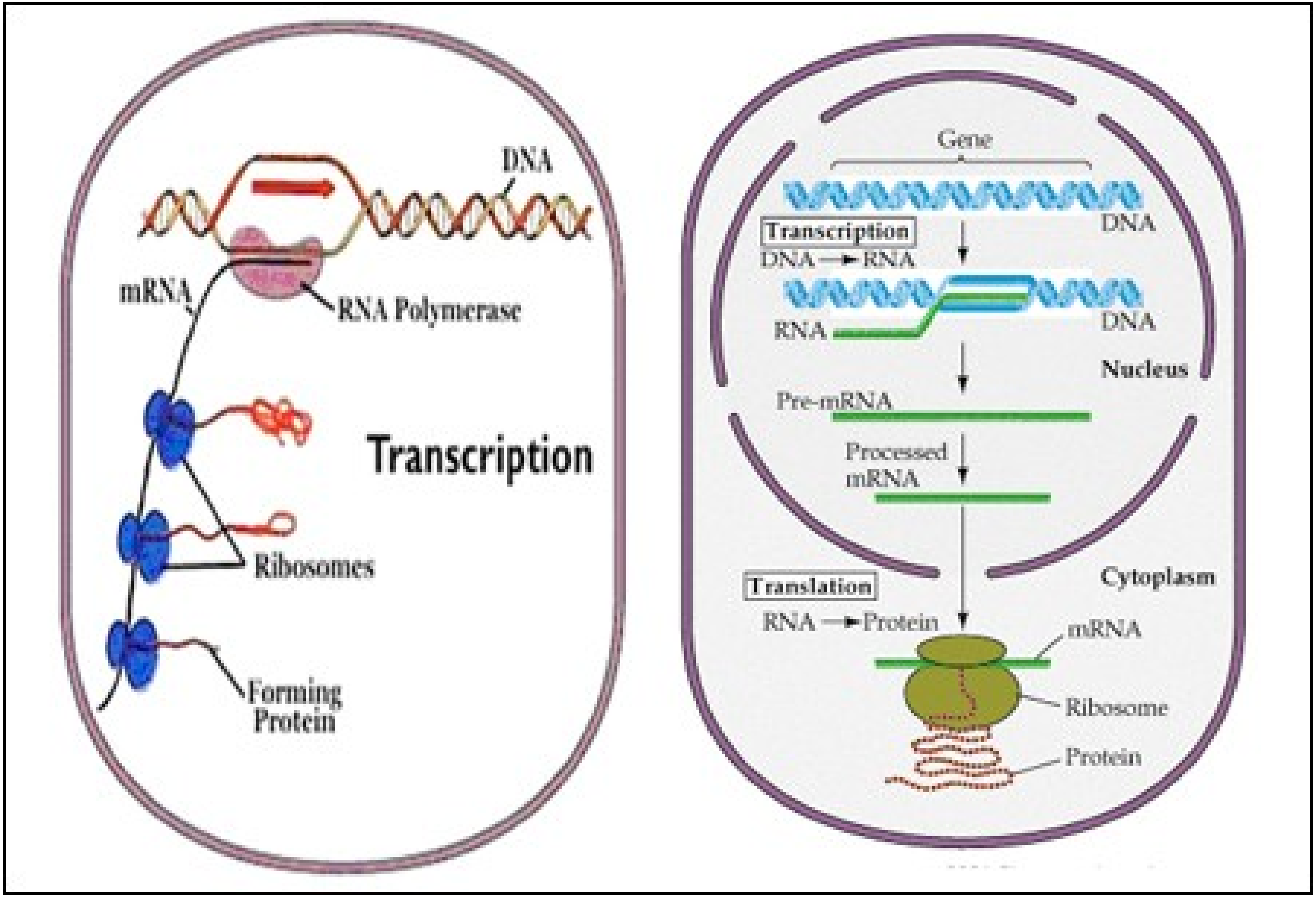
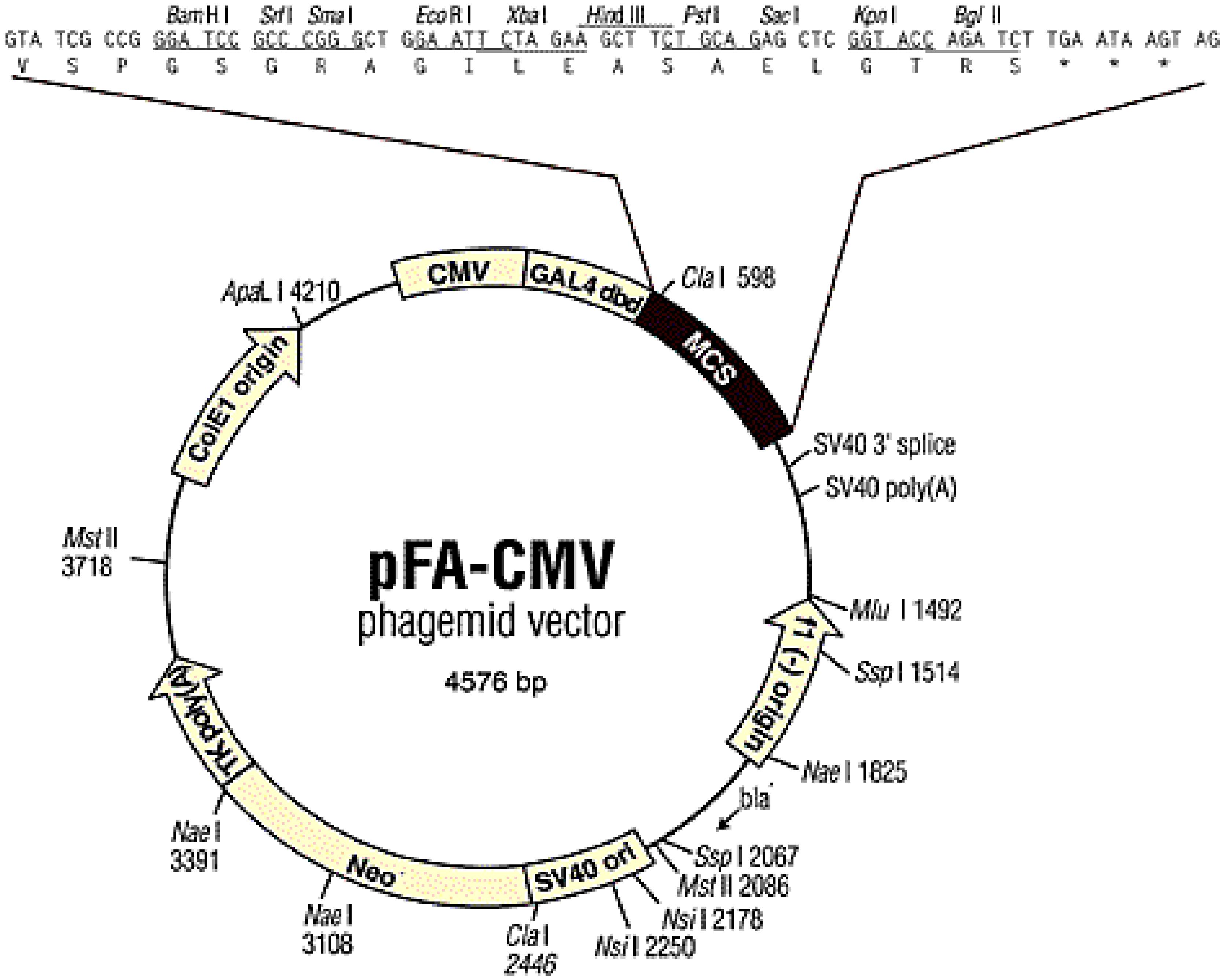
0

1

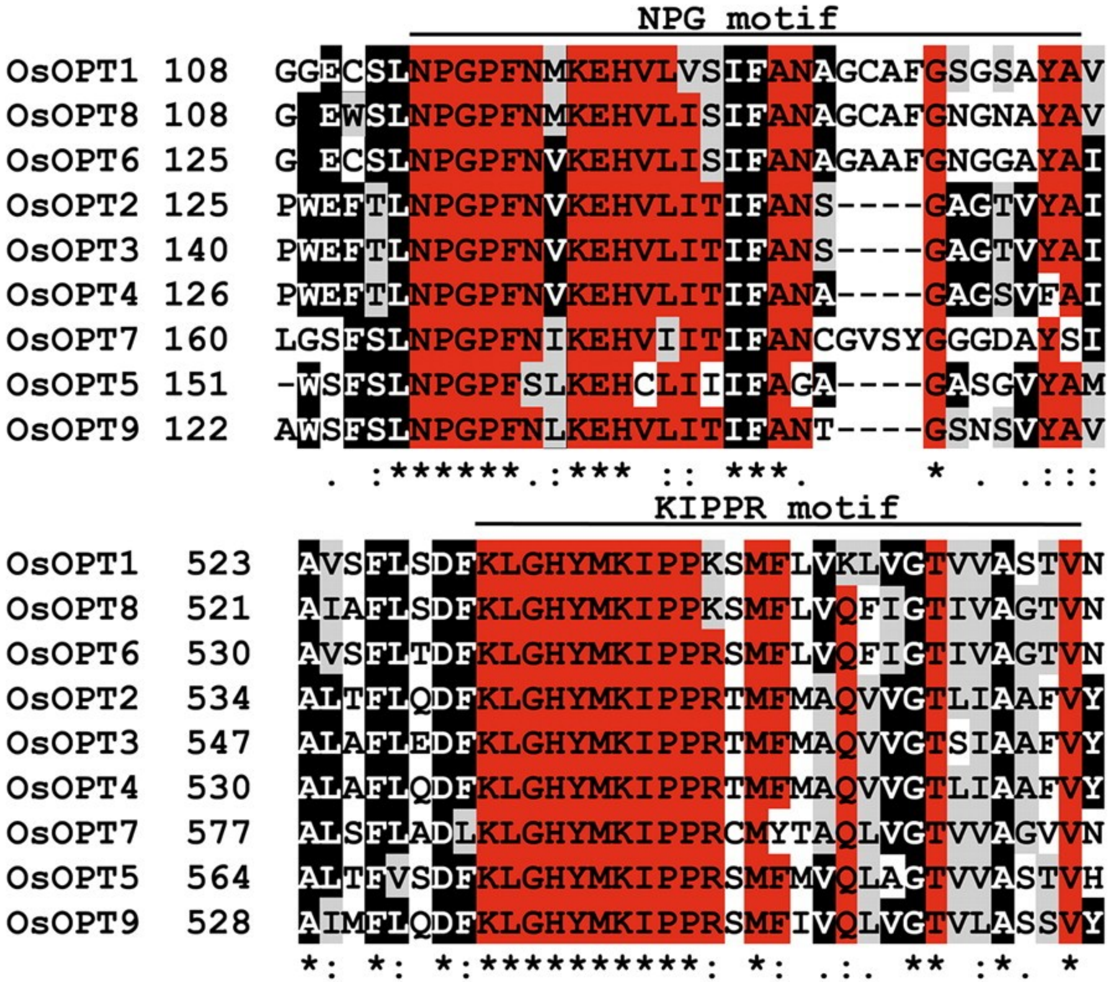

DNA
mRNA
RNA Polymerase
Transcription
Ribosomes
Forming Protein
Gene
DNA
Transcription
DNA → RNA
RNA
DNA
Nucleus
Pre-mRNA
Processed mRNA
Cytoplasm
Translation
RNA → Protein
mRNA
Ribosome
Protein

2

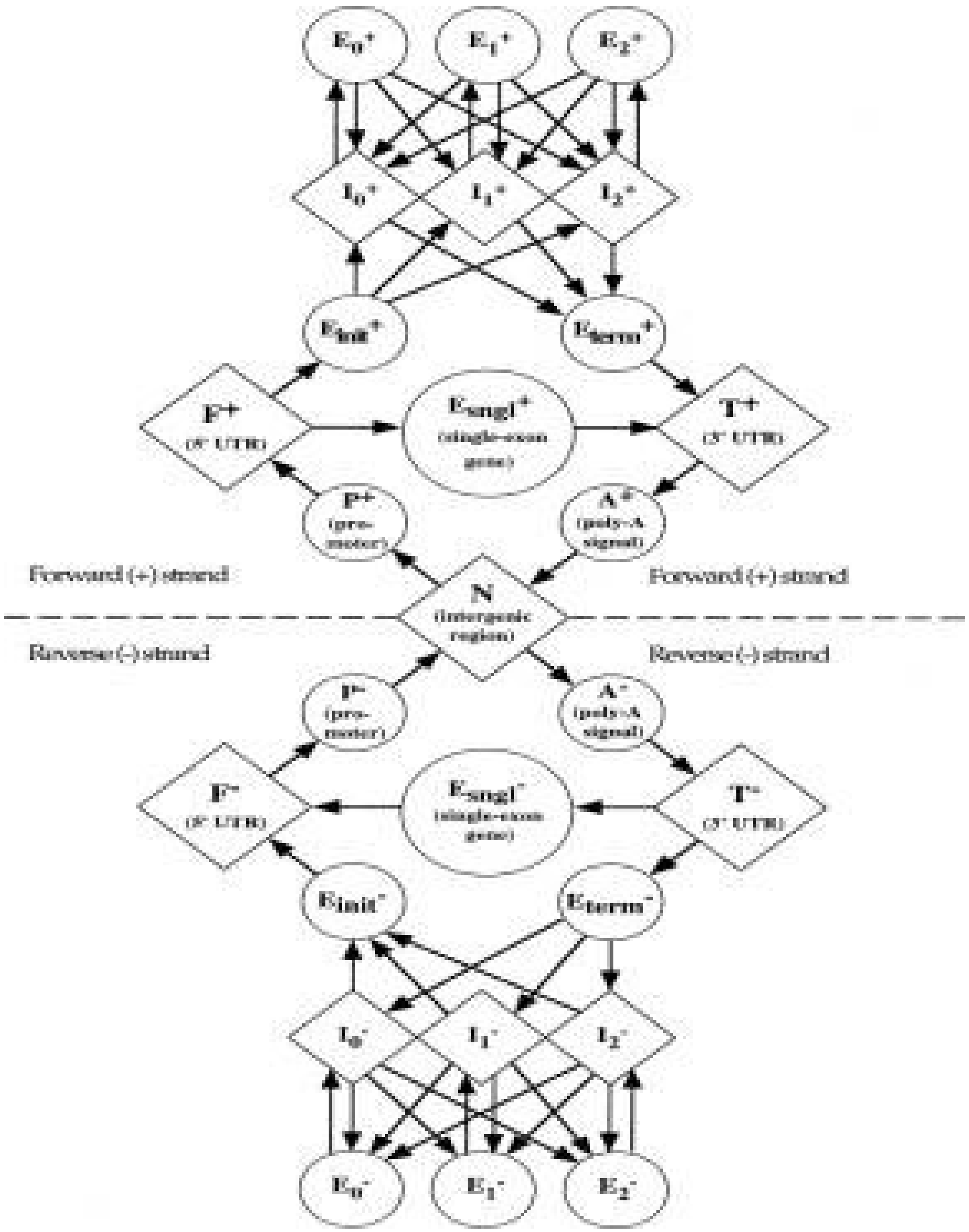

*Bam*H I
Srf I
Sma I
*Eco*R I
Xba I
*Hin*d III
Pst I
Sac I
Kpn I
Bgl II
GTA TCG CCG GGA TCC GCC CGG GCT GGA ATT CTA GAA GCT TCT GCA GAG CTC GGT ACC AGA TCT TGA ATA AGT AG
V S P G S G R A G I L E A S A E L G T R S * * *
CMV
GAL4 dbd
MCS
Cla I 598
SV40 3' splice
SV40 poly(A)
Mlu I 1492
Ssp I 1514
f1 (-) origin
Nae I 1825
bla
Ssp I 2067
Mst II 2086
Nsi I 2178
Nsi I 2250
SV40 ori
Cla I 2446
Neo
Nae I 3108
Nae I 3391
TK poly(A)
Mst II 3718
ColE1 origin
ApaL I 4210
pFA-CMV
phagemid vector
4576 bp

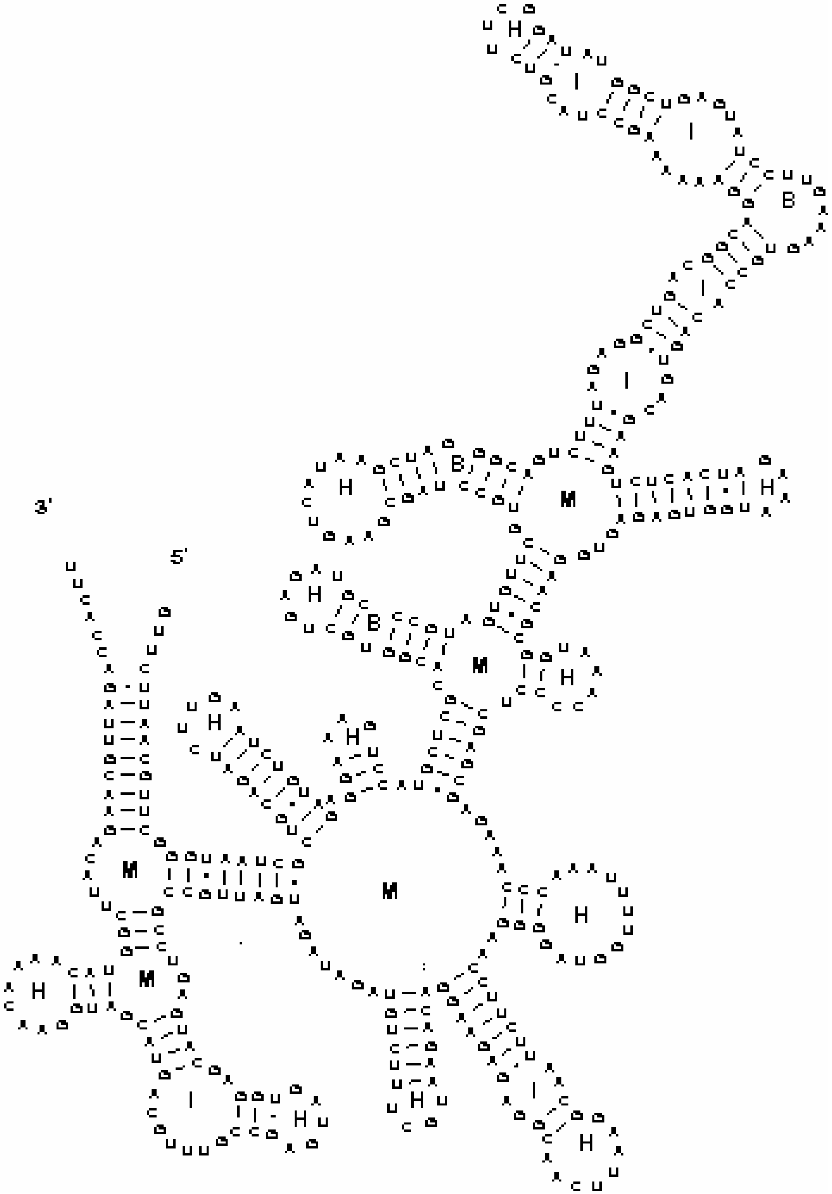
```
                                                                     Smallest
                                                                       Sum
                                                                High  Probability
Sequences producing High-scoring Segment Pairs:                 Score  P(N)      N

...............................................................................
sp|P08506|DACC_ECOLI PENICILLIN-BINDING PROTEIN 6 PRECURS...   894  5.0e-120  1
sp|P38422|DACF_BACSU PENICILLIN-BINDING DACF PROTEIN PREC...   209  5.0e-47   3
...............................................................................
sp|P28271|IREB_MOUSE IRON-RESPONSIVE ELEMENT BINDING PROT...    59  0.9996    1
sp|P31571|CAIA_ECOLI PROBABLE CARNITINE OPERON OXIDOREDUC...    48  0.9998    2

D. Alignments

..............................................................................

>sp|P08506|DACC_ECOLI PENICILLIN-BINDING PROTEIN 6 PRECURSOR
            (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE FRACTION C) (EC 3.4.16.4)
            (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (PBP-6).
            Length = 400

  Score = 894 (409.5 bits), Expect = 5.0e-120, P = 5.0e-120
 Identities = 169/342 (49%), Positives = 237/3 42 (69%)

Query:     1 MDYTTGQILTAGNEHQQRNPASLTKLMTGYVVDRAIDSHRITPDDIVTVGRDAWAKDNPV 60
             MDY +G++L  GN  ++ +PASLTK+MT YVV +A+ + +I   D+VTVG+DAWA  NP
Sbjct:    45 MDYASGKVLAEGNADEKLDPASLTKIMTSYVVGQALKADKIKLTDMVTVGKDAWATGNPA 104

Query:    61 FVGSSLMFLKEGDRVSVRDLSRGLIVDSGNDACVALADYIAGGQRQFVEMMNNYAEKLHL 120
               GSS+MFLK GD+VSV DL++G+I+ SGNDAC+ALADY+AG Q  F+ +MN YA+KL L
Sbjct:   105 LRGSSVMFLKPGDQVSVADLNKGVIIQSGNDACIALADYVAGSQESFIGLMNGYAKKLGL 164
```

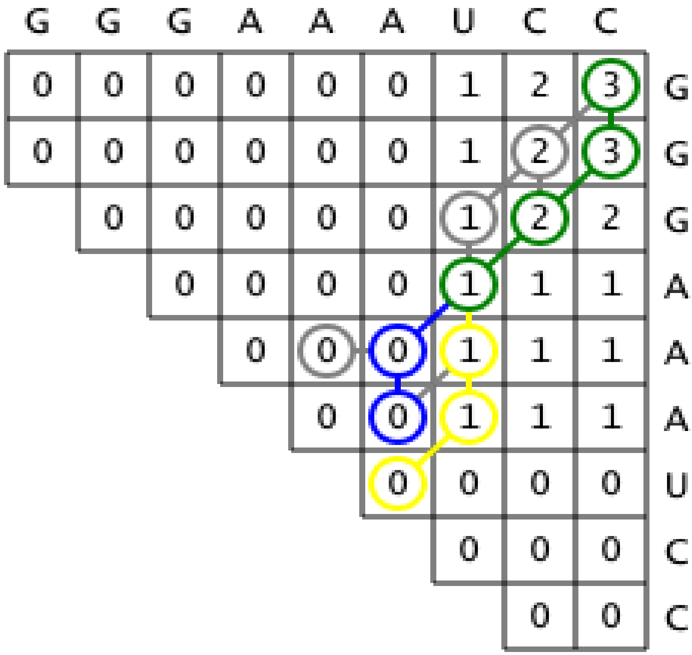
4



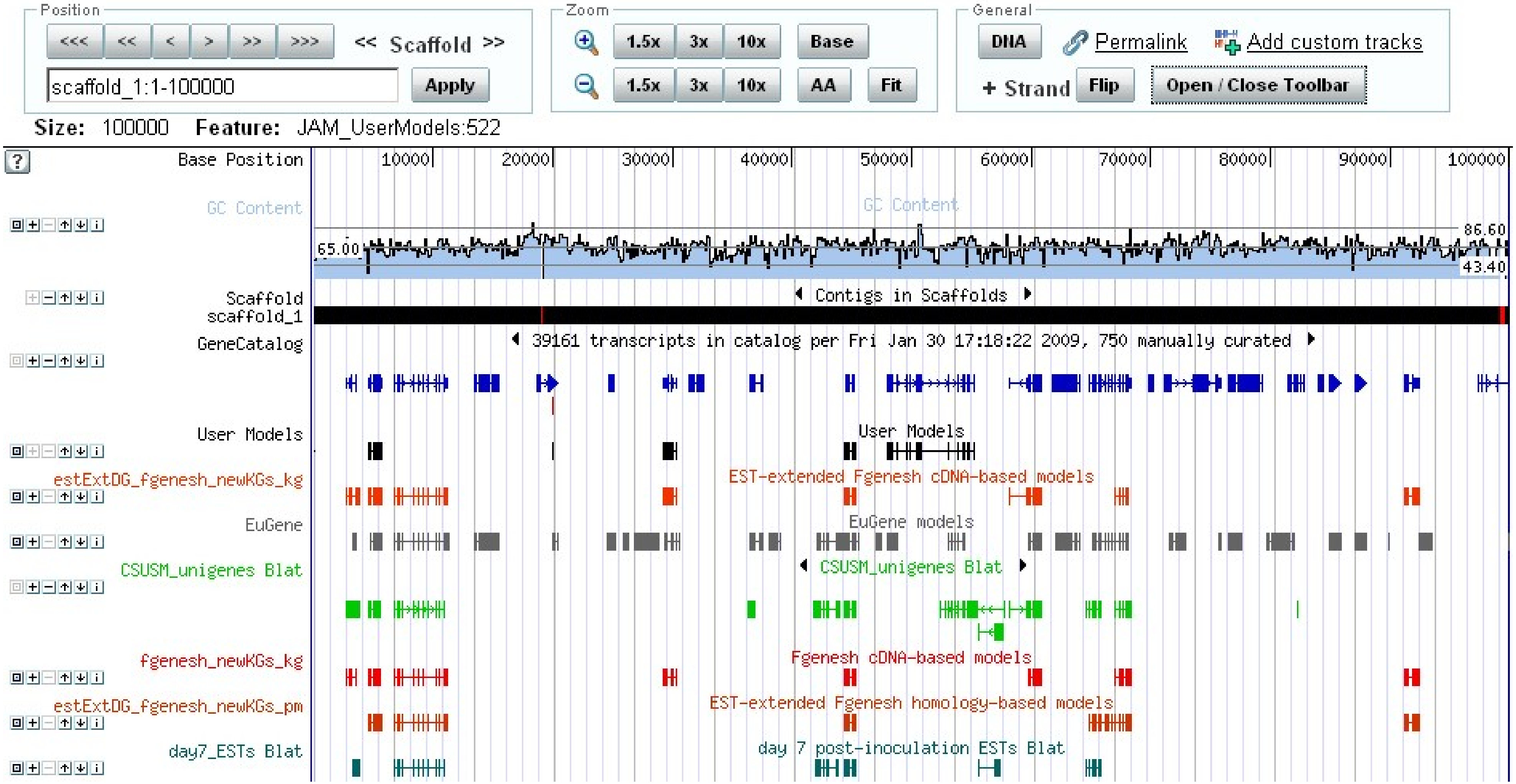
```
OsOPT1 108   GGECSLNPGPFNMKEHVLVSIFANAGCAFGSGSAYAV
OsOPT8 108   G EWSLNPGPFNMKEHVLISIFANAGCAFGNGNAYAV
OsOPT6 125   G ECSLNPGPFNVKEHVLISIFANAGAAFGNGGAYAI
OsOPT2 125   PWEFTLNPGPFNVKEHVLITIFANS----GAGTVYAI
OsOPT3 140   PWEFTLNPGPFNVKEHVLITIFANS----GAGTVYAI
OsOPT4 126   PWEFTLNPGPFNVKEHVLITIFANA----GAGSVFAI
OsOPT7 160   LGSFSLNPGPFNIKEHVIITIFANCGVSYGGGDAYSI
OsOPT5 151   -WSFSLNPGPFSLKEHCLIIIFAGA----GASGVYAM
OsOPT9 122   AWSFSLNPGPFNLKEHVLITIFANT----GSNSVYAV
               . :******.:***  ::***.     * . .:::

OsOPT1  523   AVSFLSDFKLGHYMKIPPKSMFLVKLVGTVVASTVN
OsOPT8  521   AIAFLSDFKLGHYMKIPPKSMFLVQFIGTIVAGTVN
OsOPT6  530   AVSFLTDFKLGHYMKIPPRSMFLVQFIGTIVAGTVN
OsOPT2  534   ALTFLQDFKLGHYMKIPPRTMFMAQVVGTLIAAFVY
OsOPT3  547   ALAFLEDFKLGHYMKIPPRTMFMAQVVGTSIAAFVY
OsOPT4  530   ALAFLQDFKLGHYMKIPPRTMFMAQVVGTLIAAFVY
OsOPT7  577   ALSFLADLKLGHYMKIPPRCMYTAQLVGTVVAGVVN
OsOPT5  564   ALTFVSDFKLGHYMKIPPRSMFMVQLAGTVVASTVH
OsOPT9  528   AIMFLQDFKLGHYMKIPPRSMFIVQLVGTVLASSVY
              *: *: *:**********: *: .:. ** :*. *
```


E_0^+
E_1^+
E_2^+
I_0^+
I_1^+
I_2^+
E_{init}^+
E_{term}^+
F^+ (5' UTR)
E_{sngl}^+ (single-exon gene)
T^+ (3' UTR)
P^+ (pro-moter)
A^+ (poly-A signal)
Forward (+) strand
Forward (+) strand
N (intergenic region)
Reverse (-) strand
Reverse (-) strand
P^- (pro-moter)
A^- (poly-A signal)
F^- (5' UTR)
E_{sngl}^- (single-exon gene)
T^- (3' UTR)
E_{init}^-
E_{term}^-
I_0^-
I_1^-
I_2^-
E_0^-
E_1^-
E_2^-





G	G	G	A	A	A	U	C	C	
0	0	0	0	0	0	1	2	3	G
0	0	0	0	0	0	1	2	3	G
	0	0	0	0	0	1	2	2	G
		0	0	0	0	1	1	1	A
			0	0	0	1	1	1	A
				0	0	1	1	1	A
					0	0	0	0	U
						0	0	0	C
							0	0	C

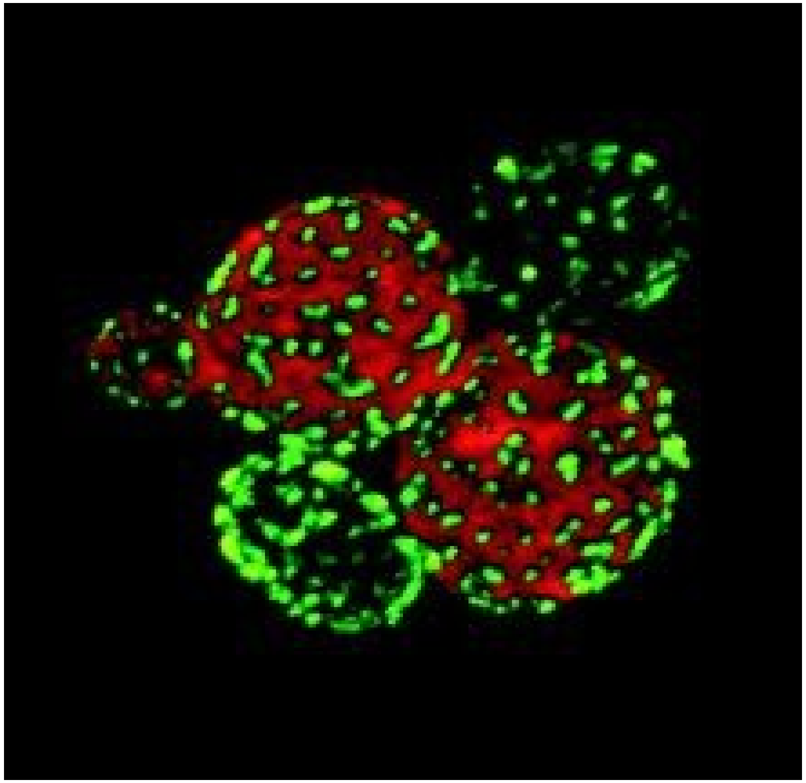


scaffold_1:1-100000

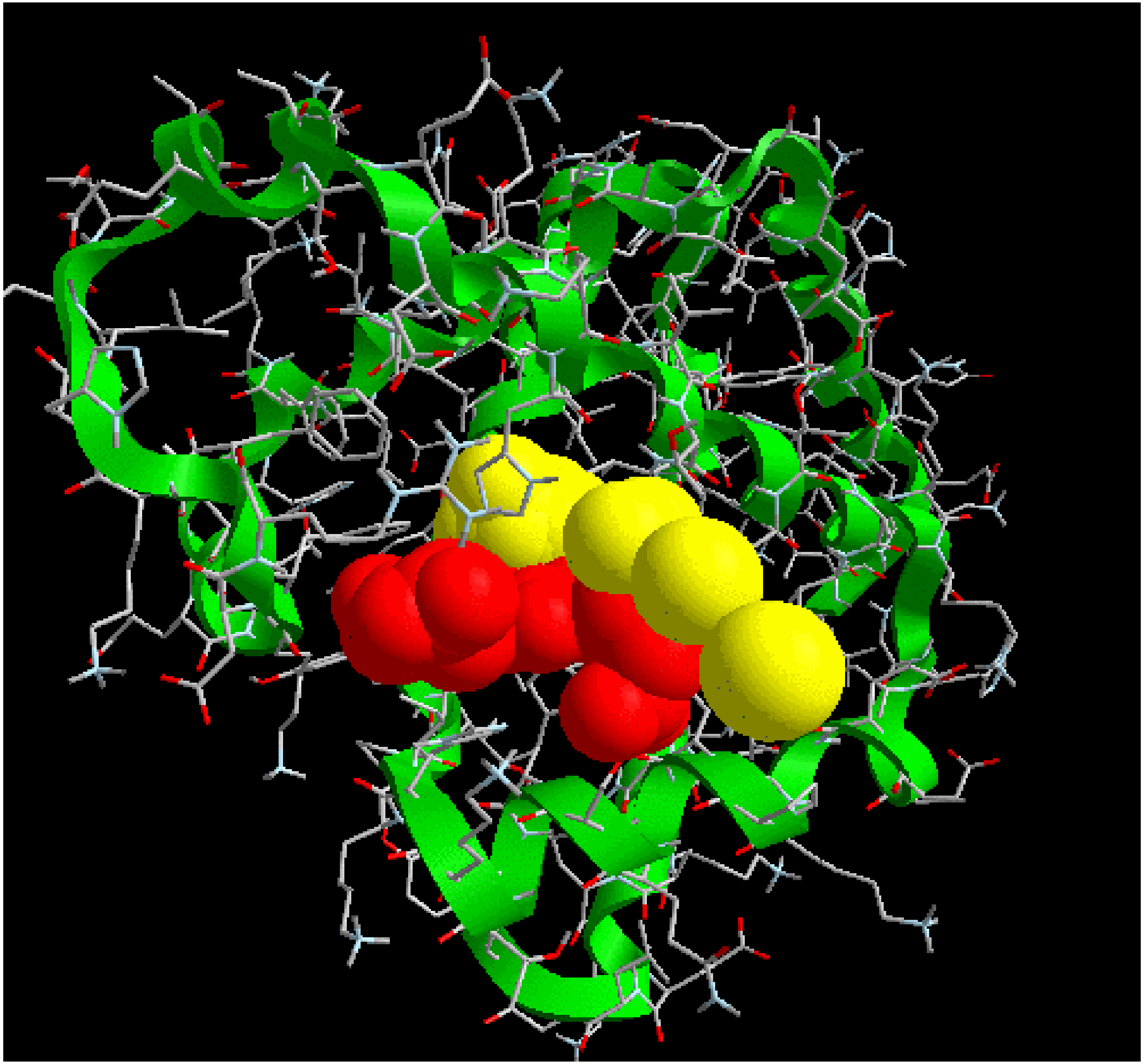
Size: 100000 **Feature:** JAM_UserModels:522

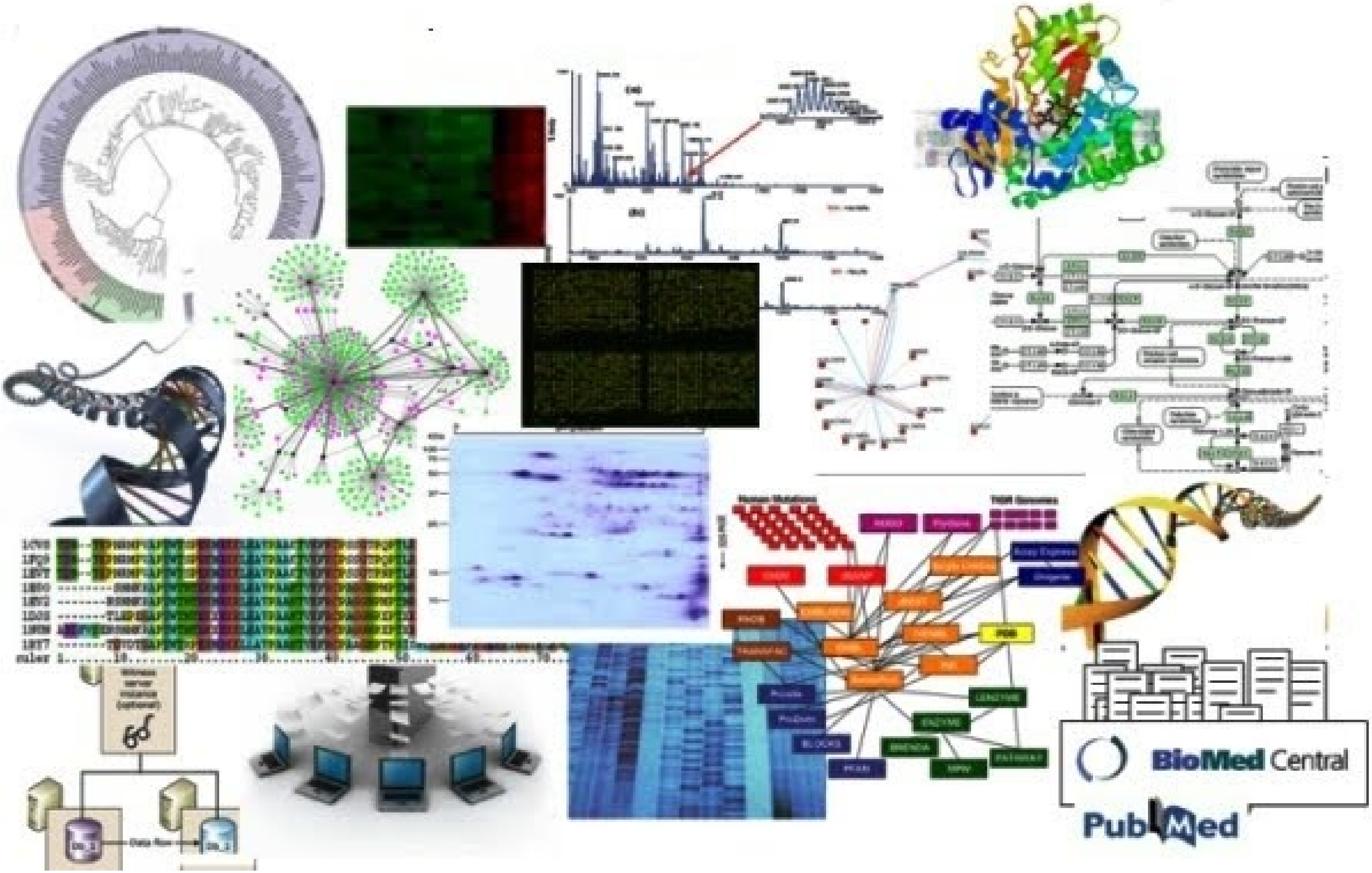
39161 transcripts in catalog per Fri Jan 30 17:18:22 2009, 750 manually curated


Science
Fo
F1
125
Science
AAAS



8




BioMed Central
PubMed

EBI - Hinxton

Sanger Sequencing Centre