GENSCANW output for sequence 21:52:54

```
GENSCAN 1.0     Date run: 25-Jun-107    Time: 21:52:58

Sequence 21:52:54 : 23767 bp : 52.07% C+G : Isochore 3 (51 - 57 C+G%)

Parameter matrix: HumanIso.smat

Predicted genes/exons:


Gn.Ex Type S .Begin ...End .Len Fr Ph I/Ac Do/T CodRg P.... Tscr..
----- ---- - ------ ------ ---- -- -- ---- ---- ----- ----- ------

 1.01 Init +     25    350  326  0  2   83   64   486 0.490  40.51
 1.02 Intr +    926   1013   88  2  1   69   55    69 0.401   2.17
 1.03 Intr +   8276   8334   59  1  2  132   48     9 0.161  -0.53
 1.04 Intr +  10089  10288  200  0  2   87   94   346 0.189  34.62
 1.05 Intr +  14776  14962  187  2  1  100   55   166 0.860  13.77
 1.06 Intr +  15420  15661  242  0  2   59  127   248 0.885  23.53
 1.07 Intr +  17518  17669  152  2  2   79  105   137 0.999  14.89
 1.08 Intr +  17797  17998  202  0  1   97  111   187 0.827  21.28
 1.09 Intr +  18730  18921  192  2  0  110  101   245 0.991  27.98
 1.10 Intr +  20524  20645  122  2  2   65  110   269 0.999  27.52
 1.11 Intr +  20739  20883  145  2  1   95   94   126 0.901  14.27
 1.12 Term +  22321  22448  128  2  2   63   41   162 0.970   7.75
```

Click here to view a PDF image of the predicted gene(s)

Click here for a PostScript image of the predicted gene(s)

```
Predicted peptide sequence(s):


>21:52:54|GENSCAN_predicted_peptide_1|680_aa
MVSKLSQLQTELLAALLESGLSKEALIQALGEPGPYLLAGEGPLDKGESCGGGRGELAEL
PNGLGETRGSEDETDDDGEDFTPPILKELENLSPEEAAHQKAVVETLLQHPTSPAGAIRG
CPFYIPIRWRTPRLLENWGDPQQRHDSQVASGPFESLWEDPWRVAKMVKSYLQQHNIPQR
EVVDTTGLNQSHLSQHLNKGTPMKTQKRAALYTWYVRKQREVAQQFTHAGQGGLIEEPTG
DELPTKKGRRNRFKWGPASQQILFQAYERQKNPSKEERETLVEECNRAECIQRGVSPSQA
QGLGSNLVTEVRVYNWFANRRKEEAFRHKLAMDTYSGPPPGPGPGPALPAHSSPGLPPPA
LSPSKVHGVRYGQPATSETAEVPSSSGGPLVTVSTPLHQVSPTGLEPSHSLLSTEAKLVS
AAGGPLPPVSTLTALHSLEQTSPGLNQQPQNLIMASLPGVMTIGPGEPASLGPTFTNTGA
STLVIGLASTQAQSVPVINSMGSSLTTLQPVQFSQPLHPSYQQPLMPPVQSHVTQSPFMA
TMAQLQSPHALYSHKPEVAQYTHTGLLPQTMLITDTTNLSALASLTPTKQVFTSDTEASS
ESGLHTPASQATTLHVPSQDPASIQHLQPAHRLSASPTVSSSSLVLYQSSDSSNGQSHLL
PSNHSVIETFISTQMASSSQ


Explanation

Gn.Ex : gene number, exon number (for reference)
Type  : Init = Initial exon (ATG to 5' splice site)
        Intr = Internal exon (3' splice site to 5' splice site)
        Term = Terminal exon (3' splice site to stop codon)
        Sngl = Single-exon gene (ATG to stop)
        Prom = Promoter (TATA box / initation site)
        PlyA = poly-A signal (consensus: AATAAA)
S     : DNA strand (+ = input strand; - = opposite strand)
```

Begin : beginning of exon or signal (numbered on input strand)
End : end point of exon or signal (numbered on input strand)
Len : length of exon or signal (bp)
Fr : reading frame (a forward strand codon ending at x has frame x mod 3)
Ph : net phase of exon (exon length modulo 3)
I/Ac : initiation signal or 3' splice site score (tenth bit units)
Do/T : 5' splice site or termination signal score (tenth bit units)
CodRg : coding region score (tenth bit units)
P : probability of exon (sum over all parses containing exon)
Tscr : exon score (depends on length, I/Ac, Do/T and CodRg scores)

Comments

The SCORE of a predicted feature (e.g., exon or splice site) is a log-odds measure of the quality of the feature based on local sequence properties. For example, a predicted 5' splice site with score > 100 is strong; 50-100 is moderate; 0-50 is weak; and below 0 is poor (more than likely not a real donor site).

The PROBABILITY of a predicted exon is the estimated probability under GENSCAN's model of genomic sequence structure that the exon is correct. This probability depends in general on global as well as local sequence properties, e.g., it depends on how well the exon fits with neighboring exons. It has been shown that predicted exons with higher probabilities are more likely to be correct than those with lower probabilities.