Fex

Prediction of internal, 5'- and 3'- exons in Human DNA sequences.

Method description:
Algorithm first predicts all internal exons in a given sequence by linear discriminant function combining characteristics describing donor and acceptor splice sites, 5'- and 3'-intron regions and also coding regions for each open reading frame flanked by GT and AG base pairs. Potential 5'- and 3'- exons are predicted by corresponding discriminant functions on the left side of the first internal exon and on the right side from last internal exon, respectively.

Accuracy:
The accuracy of precise exon recognition on the set of 210 genes (with 761 internal exons) is 70% with a specificity of 63%. The recognition quality computed at the level of individual nucleotides is 87% for exons sequences (Sp=82%) with the level 97% for intron sequences. This program does not assemble the exons and is more reliable for a case of missing exons - for example, due to sequencing errors.

Fex output:
First line - name of your sequence
Next lines - positions of predicted exons, their 'weights', ORF number and potential number ORFs for a particular exon.

For example:
```
Seq name: Adh_and_cact.1 (2919020 bases) 848501 853000
 Length of sequence:  4500  Exon thr-  0 Overlap thr-    0.0
 # of potential exons: 9
   2758 -    2936 + w=   27.96 ORF= 0 First    exon   2758 -    2934
   3291 -    3354 - w=   13.63 ORF= 2 First    exon   3292 -    3354
   2577 -    2690 + w=   11.78 ORF= 2 Internal exon   2579 -    2689
      3 -     269 + w=   10.06 ORF= 0 Single   exon      3 -     269
   3024 -    3107 - w=    9.15 ORF= 2 Internal exon   3025 -    3105
    385 -     543 + w=    2.22 ORF= 0 Last     exon    385 -     543
   3169 -    3173 + w=    2.18 ORF= 0 First    exon   3169 -    3171
   2213 -    2380 + w=    1.65 ORF= 0 Last     exon   2213 -    2380
   1037 -    1076 + w=    0.25 ORF= 0 First    exon   1037 -    1075
>Exon-     1 Amino acid sequence -    59 aa, chain +
MANCPHTIGVEFGTRIIEVDDKKIKLQIWDTAGQERFRAVTRSYYRGAAGALMVYDITR
>Exon-     2 Amino acid sequence -    21 aa, chain -
MACAELRTRRRSDRADPPGCS
>Exon-     3 Amino acid sequence -    37 aa, chain +
PNMTAAPYNYNYIFKYIIIGDMGVGKSCLLHQFTEKK
>Exon-     4 Amino acid sequence -    88 aa, chain +
MLVQTPGISKSWMSSICLRESTFFMSCDRFRRSVSHCEGDTHELTAWQRVYLATHIWHRL
AGAQVVDLHIVNFVYEHLEGRFLLKIKT
>Exon-     5 Amino acid sequence -    27 aa, chain -
NLPSALQIRFVANEKDHSAGIGEIASV
>Exon-     6 Amino acid sequence -    52 aa, chain +
CDRRKPSKTRERKSSEKRLLICIDLPIENNRNNCLSVQPRNPAKPVCVLARK
>Exon-     7 Amino acid sequence -     1 aa, chain +
M
>Exon-     8 Amino acid sequence -    55 aa, chain +
LAGKQTRSAVQTQAGLKKKYRGQFEKGEQNVVSTQNKLMQRLGLLISSDYGWTFK
>Exon-     9 Amino acid sequence -    13 aa, chain +
MVGQKRPPLYLKI
```

References:

Solovyev V.V.,Salamov A.A., Lawrence C.B. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. (Nucl.Acids Res.,1994,22,24,5156-5163).

Solovyev V.V., Salamov A.A. , Lawrence C.B. The prediction of human exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. in: The Second International conference on Intelligent systems for Molecular Biology (eds. Altman R., Brutlag D., Karp R., Latrop R. and Searls D.), AAAI Press, Menlo Park, CA (1994, 354-362).