

PolyaH

Recognition of 3'-end cleavage and polyadenylation region of human mRNA precursors.

Method description:

Algorithm predicts potential position of poly-A region by linear discriminant functions combining characteristics describing various contextual features of these sites. The default LDF threshold in the server is equal 0.

Accuracy:

The accuracy has been estimated for the set of 131 poly-A regions and 1466 non-poly-A regions of human genes, having AATAAA sequence. For 86% accuracy poly-A region prediction the algorithm has 8% false predictions (Sp=50%; C=0.62). For example, with threshold 0.7 it predicts 8 of 9 poly-A sites of AD2 genome (35937 bp.) and overpredict 4 false (Compare with method of poly-A site prediction (CABIOS 1994,10,597-603), which for 8 true predicted sites gives 968 false positive sites).

PolyaH output:

First line - name of your sequence;

2nd line - Length of your sequence

Next lines - positions of predicted sites and their 'weights', Position shows the first nucleotide of the AATAAA consensus in the predicted region

For example:

```
HSG11C4A      1741 bp      DNA      PRI      21-FEB
Length of sequence-      1741
      1 potential polyA site was predicted
Pos.:      988 LDF- 4.06
```