

GeneBuilder: Prediction of gene structure in different eukaryotic organisms.

Luciano Milanesi

Istituto di Tecnologie Biomediche Avanzate - CNR (Milano)

Prediction of protein-coding genes in newly sequenced DNA becomes very important in large genome sequencing projects. In order to integrate the promoter with the gene structure information we have developed the GeneBuilder system able to predict the main functional signals (promoters, splice sites, TATA-box and Poly-A sites) and the coding regions (CDS). The program use several parameters and modules for gene structure analysis, homology search and automatic annotations. To improve the accuracy of the gene structure prediction the GeneBuilder program is also able to define new gene models by using different exon homology levels with a protein sequence selected from a list of homologous proteins. By using a Java based graphical interface the user can visualise the gene models and the sequence pattern of features predicted (<http://www.itba.mi.cnr.it/webgene>)

References

- Milanesi, L. and Rogozin, I.B. (1998) Prediction of human gene structure. In: Guide to Human Genome Computing (2nd ed.) (Ed. M.J.Bishop), Academic Press, Cambridge, 215-259.
- Milanesi, L., Muselli, M. and Arrigo, P. (1996) Hamming-clustering method for signals prediction in 5' and 3' regions of eukaryotic genes. *Comput. Applic. Biosci.*, 13, 399-404.
- Rogozin, I.B., Milanesi, L. and Kolchanov, N.A. (1996) Gene structure prediction using information on homologous protein sequence. *Comput. Applic. Biosci.*, 12, 161-170.
- Rogozin, I.B. and Milanesi, L. (1997) Analysis of donor splice sites in different eukaryotic organisms. *J. Mol. Evol.*, 45, 50-59.