

***ProLoc*: a new program that predicts sub cellular locations of proteins**

Amit Novik, Einat Hazkani and Erez Levanon

Compugen Ltd., 72 Pinchas Rosen Street, Tel-Aviv
69512, Israel

{amitn, einat, erez}@compugen.co.il

The sub cellular location of a protein plays an important role in determination of the protein's function. We have developed *ProLoc*, a program that can accurately predict the sub cellular localization of a protein solely from its amino acids sequence. *ProLoc* predicts, with high accuracy, the localization of a protein among 22 compartments comprising the cell organelles themselves and their membranes. In addition, it divides the membrane proteins into three groups: Type 1, Type 2 and integral membrane proteins.

To achieve high levels of accuracy several different approaches were applied concomitantly. Among these were distributions of the proteins length according to compartment, amino acid composition, prediction of trans-membrane regions, recognition of unique patterns that tend to be specific to a certain organelle (such as NLS), signal peptide and anchor modeling, and the use of unique domains from Pfam that are specific to a single compartment.

Testing the program on Swissprot non-redundant, the sub cellular location of well-annotated proteins that were not part of the training set was accurately predicted as the first choice among the 22 compartments in 77% of the cases, and as the second choice in 12% of the instances. When the possibilities were narrowed down to five compartments (the secretory pathway, transmembrane, nuclear, cytoplasmic and mitochondrial), exact predictions reached 90%.

The program was developed using a vertebrate dataset, but it can easily be modified to other organisms.

Selected References:

- 1) *ProLoc*: a new program that predicts sub cellular locations of proteins Amit Novik, Einat Hazkani and Erez Levanon (in preparation)
- 2) Localization of proteins to the Golgi apparatus; Sean Munro; Trends in cell biology Vol. 8 January 1998; 11-15
- 3) Machine learning approaches for the prediction of signal peptides and other protein sorting signals; H.Nielsen, S.Brunak and G.von Hejne; protein engineering Vol. 12 no. 1 January 1999; 3-10
- 4) Protein import into Mitochondria Walter Neupert; Ann. Rev. Biochem 1997, 66; 863-917
- 5) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids; Joe P.S Makkerh, Colin Dingwall and Ronald A. Laskey; Current Biology 1996, Vol. 6 no. 8; 1025-1027