# BMC Bioinformatics

Oral presentation

# N-terminal sequence-based prediction of subcellular location

Evangelia I Petsalakis*, Pantelis G Bagos, Zoi I Litou and Stavros J Hamodrakas

Address: Department of Cell Biology and Biophysics, University of Athens Panepistimiopolis, Athens 157 01, Greece.

Email: Evangelia I Petsalakis* - epetsalakis@yahoo.com

* Corresponding author

Different compartments in a cell perform diverse tasks, thus knowledge of the localization of a protein would be highly indicative of its function. Many proteins have an N-terminal sequence of approximately 20–50 residues, which is responsible for their targeting to the appropriate location and is cleaved off, after the protein has been inserted into the organelle. Such information can be used in computational methods predicting protein localization. There are several methods for prediction of protein subcellular location available on the web, with TargetP (Emanuelsson et al 2000) being the most widely used. PredSL is a software tool, which aims to classify proteins to five subcellular locations: chloroplast, thylakoid, mitochondrion, secretory pathway and other. It combines neural networks, Markov chains, HMMs and scoring matrices in order to identify a targeting sequence at the N-terminal of a protein sequence, and determine its type (chloroplast-cTP, mitochondrial-mTP, secreted-SP, thylakoidal-lTP) and the precise location of the cleavage site. PredSL was tested on a set of 732 plant protein sequences and 637 non-plant sequences and the overall accuracy was 90.4% for the plant set and 93.3% for the non-plant set. Compared to the results obtained by TargetP when tested on the same datasets, PredSL's performance was better by 1.6% for the plant sequences and slightly worse (0.2%) for the non-plant sequences. For the prediction of thylakoid proteins PredSL was tested by cross-validation and achieved 87.3% accuracy compared to 87% by LumenP. PredSL is the only method for protein subcellular localization prediction that is available by the authors as a free, stand-alone tool, or through the URL: http://bioinformatics.biol.uoa.gr/PredSL.