Comparison of different tools for the prediction of protein interaction specificity.

B.Brannetti, G.Cesareni e M. Helmer-Citterich

Centro di Bioinformatica Molecolare, Dipartimento di Biologia, Università degli Studi di Roma Tor Vergata

Protein-protein interactions play a central role in the structural and functional aspects of the cell and the elucidation of the rules determining specificity will help understanding cell physiology and pathology. Some interesting bioinformatic methods were developed over the past few years for the inference of protein interaction specificity. These include “phylogentic profiles” (Pellegrini et al., 1999), “Rosetta stone” (Marcotte et al., 1999) and “phylogenetic trees” (Pazos and Valencia, 2001). These methods can establish functional links between proteins but they do not implicate physical interactions.

The aim of this work is to analyse and compare methods used to infer the interaction specificity of protein domains. The methods compared are all based on pep-spot data (Kramer and Schneider-Mergener, 1998) and on peptide lists that bind individual domains with sizeable affinity (Cesareni et al., 1999). The chosen techniques are: the profile method (Gribskov et al., 1987), Hidden Markov Models (Baldi, 1995), positional weight matrices derived from the peptide lists associated to the given domains and the SPOT procedure (Brannetti et al., 2000; 2001) (available at http://cbm.bio.uniroma2.it/ispot) developed to infer the interaction specificity of families of protein domains. In this analysis we focusi our interest on the SH3, PDZ and WW families of protein domains. The reliability and efficiency of each method will be evaluated essentially through their ability to associate each domain with its own binding peptides and with its own natural protein interaction partners. This last step will be perfomed using the Mint database as a source of experimental interaction data (Zanzoni et al., in press).

1. Baldi P. Substitution matrices and hidden Markov models. J Comput Biol. 1995, 2:487-91.
2. Brannetti B, Via A, Cestra G, Cesareni G, Helmer-Citterich M. SH3-SPOT: an algorithm to predict preferred ligands to different members of the SH3 gene family. J Mol Biol. 2000, 298:313-28.
3. Brannetti, B., Zanzoni, A., Montecchi Palazzi, L., Cesareni, G. and Helmer-Citterich, M. (2001). iSPOT: a web tool for the analysis and recognition of protein domain specificity. Comparative and Functional Genomics, 2:314-318.
4. Cesareni G, Castagnoli L, Cestra G. Phage displayed peptide libraries. Comb Chem High Throughput Screen. 1999, 2:1-17. Review
5. Gribskov M, McLachlan AD, Eisenberg D. Profile analysis: detection of distantly related proteins. Proc Natl Acad Sci U S A. 1987, 84:4355-4358.
6. Kramer A, Schneider-Mergener J. Synthesis of peptide libraries on continuous cellulose membrane supports. Methods Mol Biol 1998, 87:25-39.
7. Marcotte, E. M., Pellegrini, M., Ng, H-L., Rice, D.W., Yeates, T. O., and Eisenberg, D. (1999) Detecting protein function and protein-protein interactions from genome sequences. Science 285, 751-753.
8. Pazos F, Valencia A. Similarity of phylogenetic trees as indicator of protein-protein interactions. Protein Eng. 2001, 14:609-614.
9. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeats TO. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc Natl Acad Sci U S A 1999, 96:4285-4288.
10. Zanzoni A, Montecchi-Palazzi L, Quondam M, Ausiello G, Helmer-Citterich M, Cesareni G. MINT: a molecular INTeraction database. Febs Letters, in press.