iSPOT: a web-based tool for the analysis and recognition of protein domain specificity

B.Brannetti, L.Montecchi-Palazzi, A.Zanzoni, G.Cesareni, M.Helmer Citterich

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

iSPOT (iSpecificity Prediction Of Target) is a web-based tool developed to infer the protein-protein interaction mediated by families of peptide recognition modules.

The SPOT software is based on:

1) the identification of all the possible residue-residue pairs that contribute to the interaction between a domain and its ligand. This is achieved through the analysis of all the available domain-ligand complexes of known structure.

2) the construction of a matrix of position specific frequencies of residue-residue interacting pairs, using the interaction data obtained from the screening of synthetic or phage displayed libraries.

The matrix is then used to infer the peptide targets, within a given protein sequence, for a list of domains of the family module of interest.

In our first approach (Brannetti et al., 2000), all the contact positions are equally weighted. We are now applying a genetic algorithm to the optimisation of coefficients to weight the different contact positions according to their different importance in specificity. The genetic algorithm (GA) is an optimisation technique, based on the computer simulation of biological evolution, used to search the best solution to a given problem in a huge number of possible candidates.

In order to apply the GA, one must define:

1. A population of n chromosomes, containing m genes (in our case, the coefficients) and representing the candidate solutions.

2. A fitness function that assigns a value to each individual, measuring the quality of the solution.

3. A series of genetic operators, applied to the initial population to generate the offsprings, such as: a) selection, used to choose the best fitting individuals to produce offspring, b) crossing over to perform recombination between a pairs of individuals to produce one or more offsprings, c) mutation to change one or more genes of the offspring.

The SPOT_web-based tool, iSPOT, is now ready to be consulted. iSPOT is presently available for two protein modules families (SH3 and PDZ domains), and for the Major Histocompatibility Complex class I molecules, MHC I.

B.Brannetti, A.Via, G.Cestra, G.Cesareni and Helmer Citterich, M. SH3-SPOT (2000): an algorithm to predict preferred ligands of different members of the SH3 gene family. J.Mol.Biol, 298, 313-328.