Active Sequences Collection (ASC) and a new strategy to identify protein functions - (session: Novel Algorithms for Bioinformatics)

Facchiano Angelo (*), Facchiano Antonio (+), Facchiano Francesco (+)

(*) Istituto di Scienze dell'Alimentazione, CNR-via Roma 52A/C 83100 Avellino, Italy ˆ angelo.facchiano@isa.cnr.it
(+) IDI, Istituto Dermopatico dell'Immacolata-via Monti di Creta 104, 00167 Roma, Italy.

We have recently published a paper (1) describing the Active Sequences Collection (ASC), a database of short sequences, peptides or protein segments, with a demonstrated biological activity. The current version of ASC consists of three sections: DORRS, a collection of active RGD-containing peptides; TRANSIT, a collection of protein regions active as substrates of transglutaminase enzymes (TGase), and BAC, a collection of short peptides with demonstrated biological activity. ASC is aimed to create a new strategy to hypothesize biological functions of a protein. Biological activity signals may be identified by analysis of protein families, and may consist of conserved segments of sequence, sequence patterns, or conserved but not contiguous amino acids. When the active region is a very short sequence, it may be difficult to found it by means of searches against large protein databases, which return large outputs, without specific notes about functional regions. ASC database collects only biologically active segments, thus the search of a protein sequence against ASC may offer advantages into the identification of potential biologically active regions. A public version of ASC database is available at the web address http://crisceb.unina2.it/ASC/

1) Facchiano AM, Facchiano A, Facchiano F. „Active Sequences Collection (ASC) database: a new tool to assign functions to protein sequences.‰ Nucleic Acids Res 2003 Jan 1;31(1):379-82