

A probabilistic analysis of peptide distribution in proteomes - (session: Comparative Genomics and Molecular Evolution)

Luca Ferraro, Vittorio Rosato, Giovanni Giuliano

Centro Ricerche Casaccia, Unita' di Biotecnologie, Roma

We analyzed 28 complete proteomes (12 archaeal, 10 bacterial and 6 eukaryal). We defined the "representation" value Cr of a peptide in a given proteome as the ratio between its measured and expected occurrences (the latter evaluated on the basis of its AA composition). A probability value $P(Cr)$ is then evaluated on the basis of the assumption of a Poissonian distribution for the Cr .

Over-represented peptides with a $P < 0.05$ in a series of proteomes were selected. BLAST analysis shows that these peptides can be considered "motifs" in the sense that they retrieve, in different proteomes, proteins of similar function.

Data will be shown also on

- (a) the dispersion of these peptides in proteomes;
- (b) the effects of evolutionary and environmental cues on their distribution;
- (c) their coincidence with previously described protein motifs (Prosite data-base).