

DISCOVERING BIOMOLECULAR MECHANISMS WITH PROTEIN SEQUENCE STUDIES: THE ANNOTATOR SOFTWARE SUITE

FRANK EISENHABER franke@bii.a-star.edu.sg

Bioinformatics Institute, 30 Biopolis Street #07-01, Matrix, Singapore 138671

Abstract

Even when it is acknowledge that biomedical sciences are still essentially experimental and they lack a predictive theory in most subfields, it is the more important to underline the few niches where theoretical/computational approaches add creatively to the biological insight. Protein sequence analysis can predict aspects of molecular and cellular function in many cases and, in this way, decisively direct follow-up experiments for the characterization of yet uncharacterized genes and the discovery of new cellular pathways.

Therefore, the analysis of gene/protein sequences is advised to become an integral part of any molecular and cellular biological research, best in the early and planning phase since this allows avoiding unnecessary experiments. At the same time, typical mutational, expression profiling or interaction screens generate dozens or hundreds of protein targets that might require in-depth sequence analysis that, for example with available WWW-tools, will take days for a single target. The ANNOTATOR software suite provides the environment to carry out all routine steps for protein sequence analysis automatically and to enable the researcher to focus her/his time on thinking over the results. The ANNOTATOR has ca. 40 academic tools for protein sequence studies and all major databases built-in together with a number of sophisticated workflows that have shown their potential in previous discoveries.

The talk will give an insight into the biological and software design concepts of the ANNOTATOR. The application discussed include the discoveries SET domain, ATGL and Ecol functions, the prediction of various posttranslational modifications from protein sequence as well as the extension of the ANNOTATOR for protein mass-spectrometry data analysis tasks.