

Inferring New Gene and Protein Functions

Christian Schoenbach, RIKEN Genomic Sciences Center

Gene and protein functions comprise multiple transcriptional regulatory functions or protein-protein interactions. To investigate the multiple functional aspects we need to improve the use of the rich a priori knowledge in biological databases and literature. First, I discuss results and problems of a semi-automated discovery pipeline for protein motifs that has been used to identify new motifs and infer functions from more than 21,000 conceptually translated mouse cDNA sequences. In the second part I focus on the development of a discovery pipeline for protein interactions in context of disease associations that includes a text information extraction module for PubMed abstracts.