High-throughput prediction of protein-protein, -RNA and -DNA interactions in disordered regions

Zhenling Peng and Lukasz Kurgan
Department of Electrical and Computer Engineering, University of Alberta, Canada

Abstract
Intrinsically disordered proteins (IDPs) and intrinsic disordered regions (IDRs) are common in nature, with eukaryotic proteins estimated to have on average about 20% of disordered residues. IDPs and IDRs play important functional roles including transcriptional regulation, translation, and signal transduction, to name just a few. Since experimental annotations of intrinsic disorder lag behind the rapidly accumulating number of proteins, dozens of computational methods were developed for the prediction of IDRs/IDPs from protein sequences [*Cell Res* 2009, 19:929-49; *Brief Bioinform* 2010, 11:225-43; *Curr Protein Pept Sci* 2012; 13:6-18]. These predictors focus on generic intrinsic disorder, without providing insights into its function. Only a few recent studies addressed prediction of protein-protein/peptide interactions in disordered regions [*PLoS Comput Biol* 2009, 5:e1000376; *Bioinformatics* 2012, 28:i75-83].

We report on our progress to develop models for computational prediction of the protein-protein, protein-DNA, and protein-RNA binding that is mediated by IDPs and IDRs. We comparatively assess predictive quality of our models, their runtime, and provide insights into these disorder driven functions in several eukaryotic organisms. We show that our models complement outputs generated by the predictors of these interactions that were developed based on structured proteins. Computation of the three predictions takes below 1 second for a single chain using a modern desktop computer, and thus our models can be applied on the whole genome scale.