Pacific Symposium on Biocomputing 12:1-3(2007)

PROTEIN INTERACTIONS AND DISEASE

MARICEL KANN

National Center for Biotechnology Information, NIH
Bethesda, MD 20894, U.S.A.

YANAY OFRAN

Department of Biochemistry & Molecular Biophysics, Columbia University
New York, NY 10032, U.S.A.

MARCO PUNTA

Department of Biochemistry & Molecular Biophysics, Columbia University
New York, NY 10032, U.S.A.

PREDRAG RADIVOJAC

School of Informatics, Indiana University
Bloomington, IN 47408, U.S.A.

In 2003, the US National Human Genome Research Institute (NHGRI) articulated grand challenges for the genomics community in which the translation of genome-based knowledge into disease understanding, diagnostics, prognostics, drug response and clinical therapy is one of the three fundamental directions ("genomics to biology," "genomics to health" and "genomics to society").[1] At the same time the National Institutes of Health (NIH) laid out a similar roadmap for biomedical sciences.[2] Both the NHGRI grand challenges and the NIH roadmap recognized bioinformatics as an integral part in the future of life sciences. While this recognition is gratifying for the bioinformatics community, its task now is to answer the challenge of making a direct impact to the medical science and benefiting human health. Innovative use of informatics in the "translation from bench to bedside" becomes a key for bioinformaticians.

In 2005, the Pacific Symposium on Biocomputing (PSB) first solicited papers related to one aspect of this challenge, protein interactions and disease, which directly addresses computational approaches in search for the molecular basis of disease. The goal of the session was to bring together scientists interested in both bioinformatics and medical sciences to present their research progress. The session generated great interest resulting in a number of high quality papers and testable hypothesis regarding the involvement of proteins in various disease pathways. This year, the papers accepted for the session on Protein Interactions and Disease at PSB 2007 follow the same trend.

The first group of papers explored structural aspects of protein-protein interactions. Kelly et al. study ABC transporter proteins which are involved in substrate transport through the membrane. By investigating intra-transporter domain interfaces they conclude that nucleotide-binding interfaces are more conserved than those of transmembrane domains. Disease-related mutations were mapped into these interfaces. Pulim et al. developed a novel threading algorithm that predicts interactions between receptors (membrane proteins) and ligands. The method was tested on cytokines, proteins implicated in intra-cellular communication and immune system response. Novel candidate interactions, which may be implicated in disease, were predicted. Kasson and Pande use molecular dynamics to address high-order molecular organization in cell membranes. A large number of molecular dynamics trajectories provided clues into structural aspects of the insertion of about 20-residue long fusion peptide into a cell membrane by a trimer hemagglutinin of the influenza virus. The authors explain effects of mutations that preserve peptide's monomeric structure but incur loss of viral infectivity.

The second group of studies focused on analysis of protein interaction networks. Sam et al. investigate molecular factors responsible for the diseases with different causes but similar phenotypes and postulate that some are related to breakdowns in the shared protein-protein interaction networks. A statistical method is proposed to identify protein networks shared by diseases. Sridhar et al. developed an efficient algorithm for perturbing metabolic networks in order to stop the production of target compounds, while minimizing unwanted effects. The algorithm is aimed at drug development where toxicity of the drug should be reduced. Borgwardt et al. were interested in predicting clinical outcome by combining microarray and protein-protein interaction data. They use graph kernels as a measure of similarity between graphs and develop methods to improve their scalability to large graphs. Support vector machines were used to predict disease outcome. Gonzalez et al. extracted a large number of gene-disease relationships by parsing literature and mapping them to the known protein-protein interaction networks. They propose a method for ranking proteins for their involvement in disease. The method was tested on atherosclerosis. Valdivia-Granda et al. devised a method to integrate protein-protein interaction data along with other genomic annotation features with microarray data. They applied it to microarray data from a study of non-human primates infected with variola and identified early infection biomarkers. The study was complemented with a comparative protein domain analysis between host and pathogen. This work contributes to the understanding of the mechanisms of infectivity, disease and suggests potential therapeutic targets. Finally, Cook et al. worked on the novel ontology of biochemical pathways. They present Chalkboard, a tool for

building and visualizing biochemical pathways. Chalkboard can be used interactively and is capable of making inferences.

Acknowledgements

The session co-chairs would like to thank numerous reviewers for their help in selecting the best papers among many excellent submissions.

References

1. Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. *Nature* 2003; 422(6934):835.
2. Zerhouni E. The NIH roadmap. *Science* 2003; 302(5642):63.