S8-3

Functional Classes of Protein-Protein Interaction Using Correlated mRNA Expression

Cheol-Goo Hur[1*†], Suji Jo[2†], Sung Ho Yoon[1], SoonGee Jung[1], Doheon Lee[2], Jihyun Kim[1], Hyun Ah Kang[1] and Hwan Gue Cho[3]
[1]Korea Research Institute of Bioscience and Biotechnology,
[2]Department of Biosystems, Korea Advanced Institute of Science and Technology,
[3]Advanced Bioinformatics Laboratory School of Comp. Sci. & Eng., College of Engineering, Pusan National University

Rapid accumulation of functional data from gene expression profiling and protein-protein interaction allows us to establishing protein interaction maps. Predicting biological pathways based on these high-throughput data is of particular concern because it provides insights into the organization of cellular system and profound understanding of cellular functions, thereby [1]. In this study, we predicted glucose-metabolism pathways in *Hansenula polymorpha* by using gene expression and protein-protein interaction data. We made cDNA microarrays containing 5,864 ORFs of *H. polymorpha*, and hybridized them with samples gotten from cultures grown on various carbon sources including glucose. Protein interaction map of *H. polymorpha* was constructed using the published protein interaction map of *Saccharomyces cerevisiae* which has close phylogenetic relationship with model organism. We applied interaction-domain pair profile method to predict protein interaction. The method combines sequences similarity searches with clustering based on interaction patterns and interaction domain information [2,3]. Proteins were correlated with each other using correlated mRNA expression method [4,5] and protein interaction map. MIPS (Munich Information Center for Protein Sequences, 6) functional catalog and Gene Ontology (GO) were used to assign functional roles for proteins and to find metabolic pathways related to glucose metabolism.

The used methodologies, correlated mRNA expression and protein-protein interaction, is in the early stage of establishing protein interaction maps. Correlated mRNA expression method is a relatively inaccurate predictor of direct physical interaction and it is very sensitive to parameter choices and clustering methods during analysis [4]. However, the current efforts for enhancing the performances of such techniques envision that protein interaction map will be invaluable information for experiment design and contribute to reducing trials and errors.

† These authors contributed equally to this work

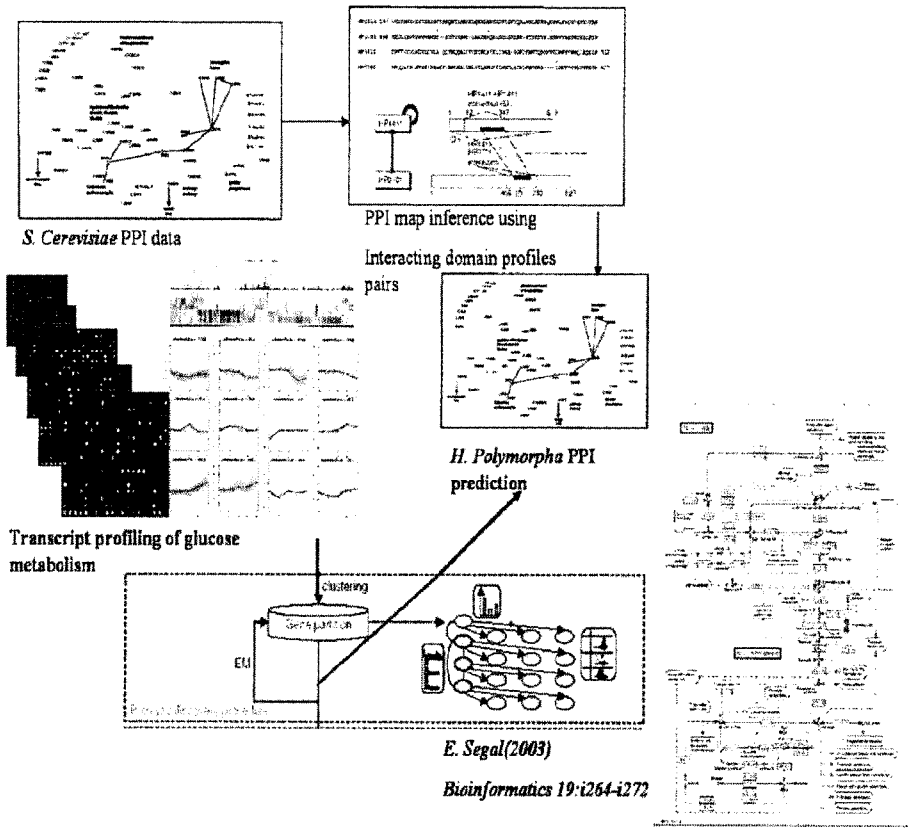

Fig 1. Schematic overview of finding out metabolic pathway based on gene expression profiling and protein-protein interaction data.

Acknowledgements

This work was supported by the 21C Frontier Microbial Genomics and Applications Center Program (grant MG02-0402-001-1-0-0) from the Ministry of Science and Technology (MOST) of the Republic of Korea.

References

1. Segal E. *et al.* (2003) Discovering molecular pathways from protein interaction and gene expression data. *Bioinformatics*. **19**:i264-i272.
2. Schwikowski B. *et al.* (2000) A network of protein-protein interactions in yeast. *Nat. Biotechnol.* **18**:1257-1261.
3. Marcotte, E.M. *et al.* (♦999) A combined algorithm for genome wide prediction of protein function. *Nature* **402**: 83 - 86.
4. von Mering C. *et al.* (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**: 399-403.
5. Ge H. *et al.* (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat. Genet.* 29: 482-486.

   Mewes H.W. *et al.* (2002) MIPS: a database for genomes and protein sequences. *Nucleic Acids Res.* 30: 31-34.