**The Organization of Human Cancer-Related Protein Complexes**

Jing Zhao1, **Sang Hoon Lee**2,3, Mikael Huss4,5, Petter Holme2,6

1Department of Mathematics, Logistical Engineering University, Chongqing, China

2IceLab, Department of Physics, Umeå University, Umeå, Sweden

3Oxford Centre for Industrial and Applied Mathematics, University of Oxford, Oxford, United Kingdom

4Science for Life Laboratory Stockholm, Solna, Sweden

5Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

6Department of Energy Science, Sungkyunkwan University, Suwon, Korea

Differential gene expression profiles for detecting disease-related genes have been intensively studied in systems biology, as a representative application of genome-wide expression data. However, it is well known that various biological functions achieved by proteins, which are encoded by genes, are consequences of different proteins’ physical binding called protein complexes. In other words, the functional units of biological functions should be protein complexes, rather than individual proteins. From this perspective, in this work, we try to switch the concept of disease-related genes in previous works to disease-related complexes, especially for 39 human solid tissue cancers and their originated normal tissues. To obtain the differential abundance levels of protein complexes, we use an optimization algorithm based on linear programming [1] applied to genome-wide differential gene expression data. From the differential abundance of complexes, we extract tissue- and cancer-specific complexes, and investigate their properties in regard to the relevance to the cancers. The method is supported by concrete biological examples presented and the clustering structures of cancer-complex bipartite relationships and those among different cancers. We believe that in the future our method and results exemplify more realistic approaches of proteomics related to diseases.

Figure: Bipartite network of cancers and protein complexes.Triangles (circles) represent cancers (complexes), respectively. The numbers (and corresponding colors) on vertices show the clustering structure defined with the Jaccard similarity index.

Reference:

[1] S. H. Lee, P.-J. Kim, and H. Jeong, BMC Syst. Biol. **5**, 126 (2011).