Northeast Structural Genomics Consortium
The Human Cancer Pathway Protein Interaction Network (HCPIN)

A comprehensive analysis of the molecular basis of cancer requires information from many fields - biochemistry, genomics, cell biology, proteomics, structural biology and systems biology. A complete understanding of the molecular interactions within cancer biology cannot be achieved without high-resolution three-dimensional structures of cancer-related proteins and their complexes. Extensive structural characterization of cancer-associated human proteins will provide key atomic details about binding interfaces and structural changes that accompany protein-protein interactions. Although multiple databases exist to help make sense of the ever-growing data on protein-protein interactions and signaling pathways, a database that integrates this information with 3D structural results has only recently been developed by the Northeast Structural Genomics Consortium (NESG) (Huang et al., 2008).

The NESG has constructed a Human Cancer Pathway Protein Interaction Network (HCPIN) by analysis of several classical cancer-associated signaling pathways and their physical protein-protein interactions, as documented in published abstracts and papers. These include pathways associated with apoptosis, cell cycle regulation, JAK/STAT, MAP Kinase, phosphotidyl inositol 3 kinase signaling, innate immune response, TGFβ signaling, and other key signaling pathways. Many well-known cancer-associated proteins, such as NfKb, TGFβ, MyD88, and EGF receptor, play central roles as "hubs" or "bottlenecks" in the HCPIN. At least half of HCPIN proteins are either directly associated with or interact with multiple signaling pathways. While some 45% of residues in these proteins are in sequence segments that meet criteria sufficient for approximate homology modeling (Blast E-val < 10^{-6}), only ~ 20% of residues in these proteins are structurally covered using high-accuracy homology modeling criteria (i.e. Blast E_val < 10^{-6} and at least 80% sequence identity) or by actual experimental structures. The HCPIN website (Fig. 1) provides a comprehensive description of this biomedical important multi-pathway network, together with experimental and homology models of HCPIN proteins useful for cancer biology research.
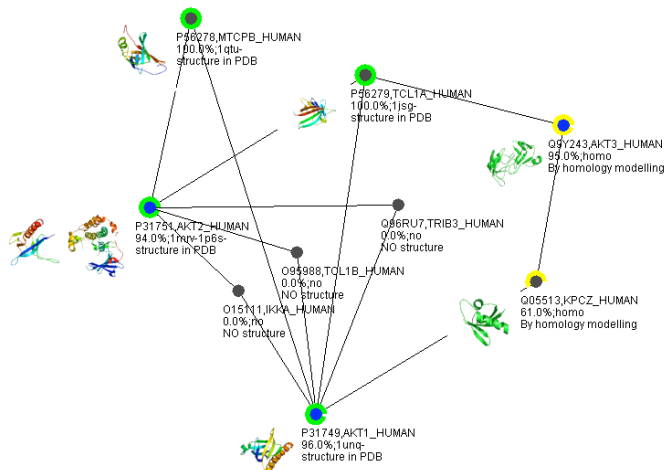

Fig 1. The Human Cancer Protein Interaction Network (HCPIN) website http://nmr.cabm.rutgers.edu/hcpin) makes structural information about cancer-associated proteins and their protein partners easily accessible. Green partial circles designate the fraction of coverage of the sequence with an experimental structure, and yellow partical circles designate the frational coverage by a high quality (> 80% seq_id) homology model.

NESG is has targeted > 1,000 human proteins (including > 2,000 protein domains) from the HCPIN for sample production and 3D structure determination [1]. Progress is also being made to determine structures of multiprotein complexes from these pathways. The long-range goal of this effort is to provide a comprehensive 3D structure-function database for human cancer-associated proteins and their complexes, in the context of their interaction networks.

The network-based target selection (BioNet) approach (Huang et al, 2007; Lu et al., 2007) is an example of a general strategy for targeting co-functioning proteins by structural genomics projects. To date, more than 35 human protein structures from the HCPIN network have been determined by X-ray crystallography and/or NMR methods. Plans for future expansion of the HCPIN database include addition of new cancer-related pathways, including metabolic pathways, and protein-protein interaction data, and incorporation of function annotation data. As part of the Community Outreach activities of NESG, we have also established collaborative interactions with cancer biologists at the Cancer Institute of New Jersey, The Ontario Cancer Institute of University of Toronto, and Herbert Irving Comprehensive Cancer Center of Columbia University, who contribute in developing the HCPIN database. We envision HCPIN as an evolving, curated resource of structure-function information for the human cancer protein interactome.

Huang, Y.J., Hang, D., Lu, L.J., Tong, L., Gerstein, M.B., and Montelione, G.T. Mol Cell Proteomics. 2008, 7, 2048-2060.Targeting the human cancer pathway protein interaction network by structural genomics.

Lu, L.J.; Sboner, A.; Huang, Y. J.; Lu, H.X.; Gianoulis; T.A.; Yip, K.Y.; Philip M. Kim, P.M.; Montelione, G.T.; Gerstein, M.B. TIBS 2007 32: 320 - 331. Comparing classical pathways and modern networks: Towards developing an edge ontology