International Journal of Advanced Engineering Research and Science (IJAERS) *[Vol-3, Issue-12, Dec- 2016]*
https://dx.doi.org/10.22161/ijaers/3.12.35 *ISSN: 2349-6495(P) | 2456-1908(O)*

Innovation of System Biological Approach in Computational Drug Discovery

Srinath Pandey[1], Ved Kumar Mishra[2], Swati Dwivedi[3], Shubhangi Dixit[4], Raghvendra Raman Mishra[5], Prashant Ankur Jain[6]

[1,2,3&4]Department of Biotechnology, Naraina Vidya Peeth Engineering and Management Institute, Naraina Group Of Institutions, Panki, Kanpur, U.P.-India.
[5]Medical Lab Technology, Institute of Medical Science, Banaras Hindu University (BHU), Varanasi
[6]Department of Computational Biology and Bioinformatics, Jacob School of Biotechnology and Bioengineering, SHIATS-Allahabad, U.P.-India.

Abstract*—Computational methods like classification and network-based algorithms can be used to understand the mode of action and the efficacy of a given compound and to help elucidating the patho-physiology of a disease. In the pharmacological industry there has already been a shift from symptomatic oriented drugs that can relieve the symptoms but not the cause of the disease to pathology-based drugs whose targets are the genes and proteins involved in the etiology of the disease. Drugs targeting the affected pathway have thus the potential to become therapeutic. A network approach to drug design would examine the effect of drugs in the context of a network of relevant protein regulatory metabolic interactions resulting in the development of a drug that would hit multiple targets selected in such a way as to decrease network integrity and so completely disrupt the functioning of the network. The screening of a compound to quickly identify the proteins it interacts with gives us all the necessary tools to identify and repair the deregulated biological pathway causing the disease.*

Keywords*— *Systems Biology, Protein Profiling, Network, Nodes & Edges, Attributes, Annotation etc.

I. INTRODUCTION

Systems Biology is the amalgamation of computer science, mathematics, physics and biology. It endeavors to study, analyze and understand complex biological systems by taking a coordinated integrated systems view using computational methodologies (1). Structure and dynamics of biological systems & prediction and inference in the complex systems is another significant implication (2).

The aim of modern systems biology is to elucidate physiology and disease from the level of molecular pathways, regulatory networks, cells, tissues, organs and ultimately the whole organism. As currently employed, the term 'systems biology' encompasses many different approaches and models for probing and understanding biological complexity, and studies of many organisms from bacteria to man (3). Because biological complexity is an exponential function of the number of system components and the interactions between them, and escalates at each additional level of organization, efforts are currently limited to simple organisms or to specific minimal pathways (and generally in very specific cell and environmental contexts) in higher organisms (4). Thus, methodologies that filter information for relevance, such as biological context and experimental knowledge of cellular and higher level system responses will be critical for successful understanding of different levels of organization in systems biology research (5). Systems biology approaches can provide a solution to design lifesaving and cost-effective drugs so that the diseases can be cured and prevented. Systems based computational techniques will be highly useful in designing effective therapeutic drugs (6, 7). System biology is defining biochemical networks in which biomolecules are showing the nodes and the molecular interactions between them are called the edges (8, 9). Analytical methods such as gene expression clustering (13), significance testing (14-17), and sequence motif identification (18) have been indispensable for enabling these discoveries and summarizing the data at each step.

II. SYSTEM BIOLOGY PATHWAY MODELING

Top-down modeling at the cell-to-organ and organism scale shows promise, but is extremely dependent on contextual cell response data. Moreover, to bridge the gap between omics and modeling, we need to collect a different type of cell biology data—data that incorporate the complexity and emergent properties of cell regulatory systems (10-12). Human tissues response can be probed *ex vivo*, an approach that, even with limitations in terms of availability and reproducibility of human tissues, has proven useful for

https://dx.doi.org/10.22161/ijaers/3.12.35 *ISSN: 2349-6495(P) | 2456-1908(O)*

validating selected compounds and targets (19). Highly reproducible or even automated approaches to cell biology, however, seem more likely to contribute to the large-scale compound and gene function analyses desired by industry and required as a basis for modeling efforts. Assays are generally designed to isolate individual pathways and to minimize biological complexity. This 'systematic biology' focus on simplified pathways is thus to be distinguished from the 'systems biology' focus on complexity and emergent properties (20,21). Complexity and emergent properties in biology derive from several features: first, complex inputs that stimulate multiple pathways; second, multiple outputs that are integrated network responses to the inputs; third, interactions between multiple cell types; and fourth, multiple contexts and environments for each cell type or combination of cell types (22). Efforts made to study cells in combination to mimic cell-cell interactions critical to *in vivo* regulatory networks and to assay cells in different complex environmental contexts (in which different combinations of pathways are activated). Parallel context or 'multisystem' analysis is important because proteins and pathways have evolved to integrate inputs and outputs from multiple contexts, so that to understand the effects of a drug (or target), data must be derived from cell responses in multiple environments (23).

A panel of just four cell systems (combinations of endothelial cells and blood mononuclear cells in four different complex inflammatory environments) was found to embody complex biology reflecting distinctive contributions of many pharmacologic targets relevant to inflammation (24)

encompassing the disease responses, profiles made up of as few as 24–40 protein readouts (including cytokines, chemokines, adhesion receptors and other inflammatory mediators) used to assess the responses of these complex systems are able to discriminate and classify most of the pathways and mechanisms effected by known modulators of inflammation, as well as a surprising array of other drugs and pathways tested. Importantly, the profiles generated from these complex, activated cell mixtures are reproducible, allowing archiving in databases and automated searching and analyses by profile similarity or other characteristics (e.g., effects on key disease-relevant parameters) (25).

For target identification and validation, informatics approaches based on the similarity of database-stored multisystem profiles have been shown to rapidly associate gene or drug activities with known (or novel) pathways, and to predict functional pathways and network interactions (26). Clustering multisystem response profiles, in which the systems are designed to capture emergent properties, can thus help define the functional architecture of signaling networks, information important (in conjunction with conventional data sets) for designing and testing computational models.

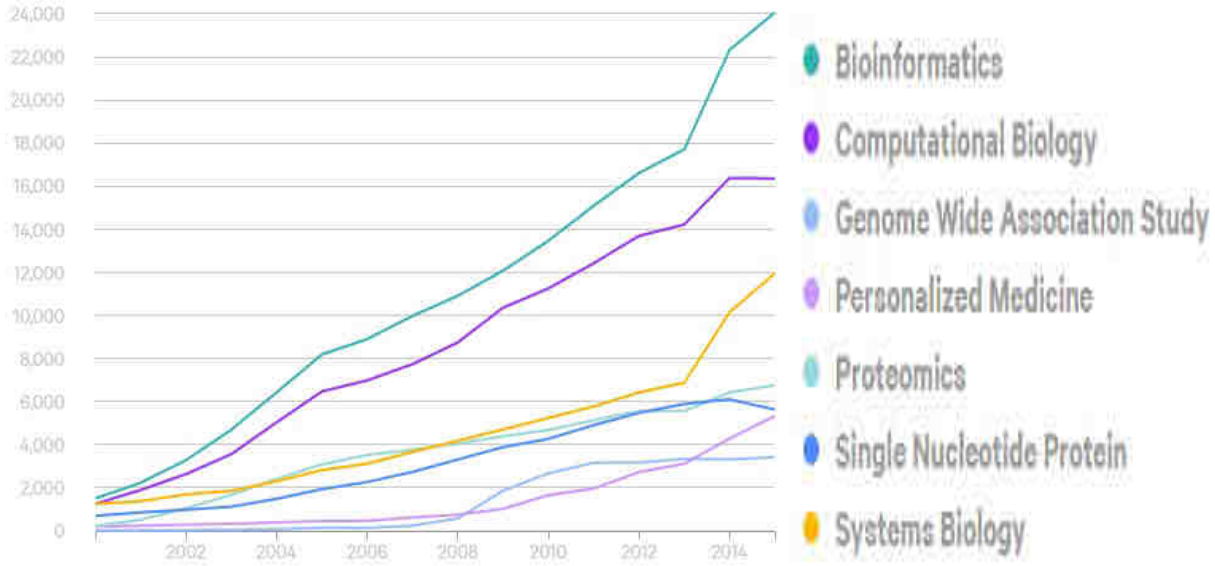

Fig.1: Research Impact from 2000 to 2015. System biology performs a search in PubMed to reveal the frequency each term appears in all publications listed in database.

https://dx.doi.org/10.22161/ijaers/3.12.35 *ISSN: 2349-6495(P) | 2456-1908(O)*

Subject / Year	Bioinformatics	Computational Biology	Genome Wide Association Study	Personalized Medicine	Proteomics	Single Nucleotide Protein	System Biology
2000	1505	1245	18	200	219	969	1246
2002	3284	2635	38	288	1040	979	1688
2004	6424	5032	71	392	2417	1486	2297
2006	8901	6984	135	462	3522	2266	3115
2008	10906	8746	572	748	4046	3310	4181
2010	13463	11255	2669	1658	4682	4271	5242
2012	16609	13690	3167	2721	5568	5474	6424
2014	22322	16380	3318	4281	6427	6105	10153
2015	24064	16359	3420	5332	6760	5628	11955

Systematic two-hybrid screens and immunoprecipitation experiments are populating the public databases with thousands of protein-protein interactions and complexes (27, 28). Other ongoing projects are defining large numbers of protein→DNA interactions (29), and protein microarrays are making it possible to map interactions between proteins and drugs, hormones, and other small molecules (30). These molecular interactions provide a paradigm for understanding changes in gene expression and for integrating a wide variety of global state measurements. One basis of systemic approaches to biological processes is the knowledge generated from genome sequencing and large-scale genetic analyses revealing an already enormous scale of interactions on the level of nucleic acids (31-33). A general understanding of the topology of genetic-interaction networks in yeast has a wider importance, because similar networks are expected to underlie the relationship between genotype and phenotype in higher eukaryotic species (34).

III. SOFTWARE FOR USE IN SYSTEM BIOLOGY

Cytoscape with its plug-ins: Cytoscape is an open source bioinformatics software tool which is used for the visualizing molecular interaction networks and integrating it with gene expression. Cytoscape was developed at the Institute of System Biology in Seattle in 2002. It is written in java and used in any java-based operating systems. Cytoscape is a project dedicated to building open-source network visualization and analysis software. Software "Core" provides basic functionality to layout and query the network and to visually integrate the network with state data. The Core is extensible through a plug-in architecture, allowing rapid development of additional computational analyses and features (35, 36). In present work we used the following four plug-ins-

BioNetBuilder: BioNetBuilder is an open-source client-server Cytoscape plug-in that offers a user-friendly interface to create biological networks integrated from several databases. Users can create networks for ~1500 organisms, including common model organisms and human. Currently supported databases include: DIP, BIND, Pro-links, KEGG, HPRD, The BioGrid, and GO, among others. The BioNetBuilder plug-in client is available as a Java Web start, providing a platform independent network interface to these public databases (37).

jActiveModulus: Identification of Modules of Seed and Neighboring Genes Using the jActiveModule Method (jAM).The MT method results in a large number of statistically significant predictions, but some of the predictions may be artifacts of low or excessive connectivity. To address this concern, the jActiveModule method is implemented to determine modules with maximal proportions of the lowest *p* value genes. jActiveModule Plug-in enumerates those ids which are actively participating in one or more function after the incorporation of compound in the microbe has already taken place. The jActiveModule plug-in gives those ids which are point of interest of our result these ids specifically are the once which depict to be actively taking post in the functioning of the microbes (38).

BiNGO: The Biological Networks Gene Ontology tool (BiNGO) is an open-source Java tool to determine which Gene Ontology (GO) terms are significantly overrepresented in a set of genes. BiNGO can be used either on a list of genes, pasted as text, or interactively on sub graphs of biological networks visualized in Cytoscape. BiNGO maps the predominant functional themes of the tested gene set on the GO hierarchy, and takes advantage of Cytoscape's versatile visualization environment to produce an intuitive and customizable visual representation of the

https://dx.doi.org/10.22161/ijaers/3.12.35 *ISSN: 2349-6495(P) | 2456-1908(O)*

results. The main advantage of BiNGO is its interactive use on molecular interaction networks, e.g. protein interaction networks or transcriptional co-regulation networks, visualized in Cytoscape. Furthermore, BiNGO offers great flexibility in the use of ontologies and annotations. Besides the traditional GO ontologies, BiNGO also supports the use of GO Slim ontologies, as well as custom ontologies and annotations (39).

MCODE: "Molecular Complex Detection" (MCODE), that detects densely connected regions in large protein-protein interaction networks that may represent molecular complexes. The method is based on vertex weighting by local neighborhood density and outward traversal from a locally dense seed protein to isolate the dense regions according to given parameters. The algorithm has the advantage over other graph clustering methods of having a directed mode that allows fine-tuning of clusters of interest without considering the rest of the network and allows examination of cluster interconnectivity, which is relevant for protein networks (40).

IV. FUTURE PROSPECTS

Targets are prioritized because they are upregulated at the gene level in disease compounds are selected to be biochemically specific); animal models are considered essential. Markup languages for gene expression data, emerging ontologies for sharing and integrating different kinds of omics and conventional biological data and the introduction of standardized high-throughput systems biology and associated informatics approaches represent important first steps on this path. Computational biology and bioinformatics approaches have the potential to completely change the way drugs are discovered and designed. A shift from symptomatic oriented drugs that can relieve the symptoms but not the cause of the disease to pathology-based drugs whose targets are the genes and proteins involved in the etiology of the disease. Drugs targeting the affected pathway have thus the potential to become therapeutic. A network approach to drug design would examine the effect of drugs in the context of a network of relevant protein-protein, regulatory and metabolic interactions resulting in the development of a drug that would hit multiple targets selected in such a way as to decrease network integrity and so completely disrupt the functioning of the network. The screening of a compound to quickly identify the proteins it interacts with gives us all the necessary tools to identify and repair the unregulated biological pathway causing the disease.

V. CONCLUSION

The systems biology approaches outlined here are contributing to meaningful drug development decisions by accelerating hypothesis-driven biology. Modeling specific physiologic problems in target validation or clinical physiology and by providing rapid characterization and interpretation of disease-relevant cell and cell system level responses. An analogy can be drawn to the genome project, in which multiple individual efforts contributed technology and informatics approaches that eventually enabled a concerted ‘big science’ push to sequence the genome. The application of systems biology to medical practice is the future of medicine. Its realization will see drug discovery and the design of multiple drug therapies and therapeutic gene circuits being pursued just as occurs now with modern, complex engineering products. The daunting task ahead is to investigate our identified sub networks in the laboratory. Because large interaction networks are suspected to contain many false-positives, an initial experiment would be to verify that the interactions in each sub network are reproducible and present under the subnet’s particular set of conditions. Routine network screening has to be performed to define novel modes of regulation, to identify evolutionarily conserved pathways, or to interrogate regulatory circuits responding to the entire spectrum of drugs and human diseases.

ACKNOWLEDGEMENT

We would like to express our deepest sense of gratitude for Dr Prashant Ankur Jain, Assistant Professor and In-charge of Department of Computational Biology and Bioinformatics, Jacob School of Biotechnology and Bioengineering, SHIATS-Allahabad and Dr Raghvendra Raman Mishra, Assistant Professor, Medical Lab Technology, Institute of Medical Science, Banaras Hindu University (BHU), Varanasi, for support of this study.

REFERENCES

[1] Lodhi H. and Muggleton S., (2004), Modeling metabolic pathways using stochastic logic programs based Ensemble methods. Second International Conference on Computational Methods in System Biology (CMSB04), LNCS, pages 119–133.

[2] Noble D., (2006), The Music of Life, Oxford University Press, pages 263-274.

[3] Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372 (2001).

[4] Ideker, T. & Lauffenburger, D. Building with a scaffold: emerging strategies for high to low-level

https://dx.doi.org/10.22161/ijaers/3.12.35 *ISSN: 2349-6495(P) | 2456-1908(O)*

cellular modeling. *Trends Biotechnol*. **21**, 255–262 (2003).

[5] Hunter, P.J. & Borg, T.K. Integration from proteins to organs: the Physiome Project.*N at. Rev. Mol. Cell Biol*. **4**, 237–243 (2003).

[6] Cho C. R., Mark Labow, Mischa Reinhardt, Jan van Oostrum and Manuel C. Peitsch (2006), The application of systems biology to drug discovery. Current Opinion in Chemical Biology, 10(4):294–302.

[7] Kumar N., Hendriks BS, Janes KA, Graaf D, Lauffenburger DA (2006), Applying computational modeling to drug discovery and development. Drug Discovery Today, 11(1718): 806–811.

[8] Materi W. and Wishart S., (2007), Computational systems biology in drug discovery and development: methods and applications. Drug Discovery Today, 12(78): 295–303.

[9] CajaMurcia F., (2006), Red Espanola de Biología de Sistemas, Systems biology, 14(1):19–31.

[10] Apic G., Ignjatovic T. and Boyer S., (2005), Illuminating drug discovery with biological pathways. FEBS Lett. 579, 1872-1877.

[11] Butcher E. C., Berg E. L., and Kunkel E. J., (2004), System biology in drug discovery. Nat Biotechnol, 22:1253-1259.

[12] Fishman M. C. and Porter J. A., (2005), A new grammar for drug discovery. Nature, 437:491-493.

[13] Eisen M. B., Spellman P. T., Brown P. O., and Botstein D. (1998). Cluster analysis and display of genome-wide expression patterns, Proc Natl Acad Sci U S A 95, 14863-8.

[14] Kerr, M. K., and Churchill, G. A. (2001). Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments, Proc Natl Acad Sci U S A 98, 8961-5.

[15] Ideker T., Thorsson V., Siegel A., and Hood L. (2000). Testing for differentially-expressed genes by maximum likelihood analysis of microarray data, J Comput Biol 7, 805-817.

[16] Rocke, D. M., and Durbin, B. (2001). A model for measurement error for gene expression arrays, J Comput Biol 8, 557-69.

[17] Pilpel, Y., Sudarsanam, P., and Church, G. M. (2001). Identifying regulatory networks by combinatorial analysis of promoter elements, Nat Genet 29, 153-9.

[18] Coleman, R.A., Bowen, W.P., Baines, I.A., Woodrooffe, A.J. & Brown, A.M. Use of human tissue in ADME and safety profiling of development candidates. *Drug Discov. Today* **6**, 1116–1126 (2001).

[19] Chanda, S.K. *et al.* Genome-scale functional profiling of the mammalian AP-1 signaling pathway. *Proc. Natl. Acad. Sci. USA* **100**, 12153–12158 (2003).

[20] Haggarty, S.J., Koeller, K.M., Wong, J.C., Butcher, R.A. & Schreiber, S.L. Multidimensional chemical genetic analysis of diversity-oriented synthesis-derived deacetylase inhibitors using cell-based assays. *Chem. Biol*. **10**, 383–396 (2003).

[21] Borisy, A.A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. USA* **100**, 7977–7982 (2003).

[22] Marton, M.J. *et al*. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* **4**, 1293–1301 (1998).

[23] Kunkel, E.J. *et al*. An integrative biology approach for analysis of drug action in models of human vascular inflammation. *FASEB J.* **18**, 1279–1281 (2004).

[24] Kunkel, E.J. *et al.* Rapid structure-activity and selectivity analysis of kinase inhibitors by BioMAP analysis in complex human primary cell-based models. *Assay Drug Dev. Technol.* **2**, 431–441 (2004).

[25] Plavec, I. *et al.* Method for analyzing signaling networks in complex cellular systems. *Proc. Natl. Acad. Sci. USA* **101**, 1223–1228 (2004).

[26] Gavin A.C., Bösche M., Krause R., Grandi P., and Marzioch M. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes, Nature 415, 141-147.

[27] Uetz, P., Giot L., Cagney G., Mansfield T. A., Judson R. S., Knight J. R., Lockshon D., Narayan V., Srinivasan M., Pochart P., Qureshi-Emili A., Li Y., Godwin B., Conover D., Kalbfleisch T., Vijayadamodar G., Yang M., Johnston M., Fields S. and Rothberg J. M., (2000), A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae.* J. M. Nature, 403, 623-627.

[28] Ren B., Robert F., Wyrick J. J., Aparicio O., Jennings E. G., Simon I., Zeitlinger J., Schreiber J., Hannett N. and Kanin E., (2000). Genome-wide location and function of DNA binding proteins, Science 290, 2306-9.

[29] Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S. and Houfek, T., (2001). Global analysis of protein activities using proteome chips, Science 293, 2101-5.

[30] Boone C., Howard B. and Brenda J. Andrews, (2007), Exploring genetic interactions and networks with yeast. Nat. Rev. Genet. 8, 437-449.

[31] Davierwala A. P., Haynes J., Li Z., Brost R. L., Robinson M. D., and Yu L., (2005), The synthetic genetic interaction spectrum of essential genes. Nat. Genet. 37, 1147-1152.
[32] Tong A. H. Y., Lesage G, Bader GD, Ding H, Xu H, Xin X, and Young J., (2004), Global mapping ofthe yeast genetic interaction network. Science, 303, 808-813.
[33] Dolinski K. and Botstein D., (2005), Changing perspectives in yeast research nearly a decade after the genome sequence. Genome Res., 15, 1611-1619.
[34] Bell G. W. and Lewitter F. (2006), Visualizing networks; Meth Enzymol. 411:408-21. Doi: 10.1016|50076-6879(06)11022-8 PMID.16939803.
[35] Shannon P., Markiel A., and Ozier O., (2003), Cytoscape: a software invironment for integrated models of biomolecular interaction networks; Genome Res. 13(11):2498-504. Doi: 10.1101|gr|239303. PMC 403769, PMID 1457658.
[36] Avila-Campillo I., Drew K., Lin J., Reiss D. J. And Bonneau R., (2007), BioNetBuilder: automatic integration of biological networks; Bioinformatics1:23(3): 392-3, EPub, PMID: 171385.
[37] Ideker T., Ozier O., Schwikowski B., and Siegel A. F., (2002), Discovering regulatory and signaling circuits in molecular interaction networks. Bioinformatics. 18 suppl 1:5233-40; PMID: 12169552.
[38] Ashburner M., Ball C. A., Blake J. A. and Botstein D., (2000), The Gene Ontology: tool for the unification of biology, Nature Genetics, 25, 25-29.
[39] Maere Steven, Heymans K. and Kuiper M., (2005). BiNGO: a cytoscape plugin to access overrepresentation of gene ontology categories in biological networks. Bioinformatics 21, 3448-3449, doi: 10.1093|bioinformatics|bti551.
[40] Gary D. B. and Christopher W. V. H., (2003), An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 4:2 doi 10, 1186|1471-2105-4-2.