Journal of Advances in Science and Technology
Vol. 12, Issue No. 24, November-2016, ISSN 2230-9659

An Analysis upon the Contribution of Protein Interaction Systems in Duseases

Rajeev Kumar[1]* Dr. Alok Mishra[2]

[1]Research Scholar, SSSUTMS, Sehore

[2]UTD, SSSUTMS, Sehore

Abstract – The study of protein-protein interactions is essential to define the molecular networks that contribute to maintain homeostasis of an organism's body functions. Disruptions in protein interaction networks have been shown to result in diseases in both humans and animals. Monogenic diseases disrupting biochemical pathways such as hereditary coagulopathies (e.g. hemophilia), provided a deep insight in the biochemical pathways of acquired coagulopathies of complex diseases. Indeed, a variety of complex liver diseases can lead to decreased synthesis of the same set of coagulation factors as in hemophilia. Similarly, more complex diseases such as different cancers have been shown to result from malfunctions of common proteins pathways. In order to discover, in high throughput, the molecular underpinnings of poorly characterized diseases, we present a statistical method to identify shared protein interaction network(s) between diseases. Integrating (i) a protein interaction network with (ii) disease to protein relationships derived from mining Gene Ontology annotations and the biomedical literature with natural language understanding (PhenoGO), we identified protein-protein interactions that were associated with pairs of diseases and calculated the statistical significance of the occurrence of interactions in the protein interaction knowledgebase.

Significant correlations between diseases and shared protein networks were identified and evaluated in this study, demonstrating the high precision of the approach and correct non-trivial predictions, signifying the potential for discovery. In conclusion, we demonstrate that the associations between diseases are directly correlated to their underlying protein-protein interaction networks, possibly providing insight into the underlying molecular mechanisms of phenotypes and biological processes disrupted in related diseases.

During a decade of proof-of-principle analysis in model organisms, protein networks have been used to further the study of molecular evolution, to gain insight into the robustness of cells to perturbation, and for assignment of new protein functions. Following these analyses, and with the recent rise of protein interaction measurements in mammals, protein networks are increasingly serving as tools to unravel the molecular basis of disease. We review promising applications of protein networks to disease in four major areas: identifying new disease genes; the study of their network properties; identifying disease-related subnetworks; and network-based disease classification.

Applications in infectious disease, personalized medicine, and pharmacology are also forthcoming as the available protein network information improves in quality and coverage.

----------♦----------

INTRODUCTION

Currently, common diseases are mainly defined by their clinical appearance, with little reference to their molecular mechanism. For example, syndromes are defined in medicine as a set of phenotypes which, occurring together, serve to define a trait or disease. These phenotypes overlap in the case of many syndromes. This overlap brought about the concept of 'syndrome families' though consideration of the commonality of features shared between diseases. Conceptually, what we have learned about 2000 human single gene diseases with a defined genetic phenotype is that each monogenic disease has a specified collection of specific phenotypic features. For example, hemophilia's with deficiencies in coagulation factors, otherwise called hereditary coagulopathies, are single gene diseases with clear Mendelian inheritance that have provided significant insight in the biochemical pathways of acquired coagulopathies. Indeed, a variety of complex liver diseases can lead to decreased synthesis of the same set of coagulation factors as in hemophilia, leading to the same disease phenotype despite very

different causes. In some cases, the clustering of syndromes into these families in combination with genetic insights has led to the discovery that what were often thought as two different disorders were really variable expressions of the same disorder. Conversely, it has long been known that mutations at different loci can lead to the same genetic disease. It has also been hypothesized that this genetic heterogeneity has its roots at the protein interaction level, suggesting that other genes associated with the phenotype also have some functional role. Therefore, it is plausible to theorize that phenotypic overlap of diseases may reflect, at multiple biological scales, the relationships and functional properties of shared underlying molecular networks. As signal transduction pathways are less understood than biochemical pathways, protein-protein interactions networks provide unique opportunities for exploring the signaling pathways of diseases.

The shift in focus to systems biology has resulted in an increased interest in biological pathways and protein-protein interaction networks. As a result, large scale knowledge bases representing them are being rapidly developed. These resources enable us to study complex biological problems using high throughput computational tools. While there is a wealth of protein-disease relationships in the published literature and a number of readily computable protein-protein interaction resources, there has been a paucity of work relating diseases using protein interactions from this kind of knowledge. Making use of these networks is a relatively new challenge in the field. Network-based analyses have been developed with a number of goals in mind, including protein function prediction, identification of functional modules, interaction prediction, and the study of network structure and evolution.

To explore the possibility of using protein-protein interaction networks to identify correlations between diseases, we hypothesize that protein-protein interactions shared by two diseases or more can be accurately identified in a protein interaction network by integrating knowledge from the literature and using statistical methods.

Research on protein-protein interaction networks and diseases has been rapidly increasing in the last two or three years. Many PPI-based methods have been proposed, each with a different way of exploiting the key assumption that “the network-neighbor of a disease gene is likely to cause the same or a similar disease”. In an early work, disease genes were uncovered by topological features in human PPI networks using the k-nearest neighbor algorithm. Because of the sparseness of other proteomic/genomic data associated with certain diseases, several PPI-based methods require the integration of heterogeneous biomedical data in order to understand the complex interplay between genes/proteins and diseases. A disease gene classification system has been proposed, to integrate the topological features of protein interaction networks with sequence and other features, and to analyze these features using support vector machines.

We have come a long way from “one-gene/one-enzyme/one function” concept originally framed by Beadle and Tatum. They provided a basic explanation of how genes work at the molecular level, however, we now know that the picture is more complex. Biological processes inside our body are governed by the well-defined organization of proteins into complexes, which perform different functions acting as molecular machines. Major biological processes such as immunity (antigen–antibody interaction), metabolism (enzyme–substrate interaction), signaling (interaction of messenger molecules, hormones, neurotransmitters with their cognate receptors), and gene expression (DNA–protein interactions), as well as the building of supramolecular assemblies (collagens, elastic fibers, actin filaments) and molecular machines (molecular motors, ribosomes, proteasome) were mediated through protein interactions.

Studying the interactome, which is the whole set of molecular physical interactions between biological entities in cells and organisms, is essential in understanding howgene functions and regulations are integrated at the level of an organism.

The notion that, a disease is rarely a consequence of an abnormality on a single gene, but it is usually the result of complex interactions and perturbations involving large sets of genes and their relationships with several cellular components, has led to the development of the network based approaches to understand human disease.

Theoretical advances in network science and paralleling advances in high-throughput efforts to map biological networks have provided a conceptual framework with which we interpret large interactome network maps. Protein–protein interaction (PPI) networks are increasingly serving as tools to decipher the molecular basis of diseases. Furthermore, the sequencing of the genome and advances in proteomics leads to the identification of proteins of unknown functions. Interaction networks might give clues on the functions of these newly discovered proteins or on new functions of already identified proteins.

The promising applications of PPI networks to disease datasets are concentrated on four major areas: (i) the identification of genes and proteins associated with diseases, (ii) the study of network properties and their relation to disease states, (iii) the

identification of disease-related subnetworks, and (iv) network-based disease classification.

Global understanding of networkswill allowresearchers to examine the disease pathways and identify strategies to control them. The integration of functional genomic and proteomic data to obtain dynamic network analysis will further improve the success of medical research.

BACKGROUND

Protein-Protein Interactions -

Protein-protein interactions are specific interactions between two or more proteins. The following is the summary of general characteristic of protein-protein interactions. Classification: Protein-protein interactions can be arbitrarily classified based on the proteins involved (structural or functional groups) or based on their physical properties (weak and transient, non-obligate vs. strong and permanent). Protein interactions are usually mediated by defined domains, hence interactions can also be classified based on the underlying domains.

Universality: All of molecular biology is about protein-protein interactions. Protein-protein interactions affect all processes in a cell: structural proteins need to interact in order to shape organelles and the whole cell, molecular machines such as ribosomes or RNA polymerases are hold together by protein-protein interactions, and the same is true for multi-subunit channels or receptors in membranes.

Specificity: distinguishes such interactions from random collisions that happen by Brownian motion in the aqeous solutions inside and outside of cells. Note that many proteins are known to interact although it remains unclear whether certain interactions have any physiological relevance.

Protein-protein interactions and protein complexes: Most protein-protein interactions are detected as interacting pairs or as components of protein complexes. Such complexes may contain dozens or even hundreds of protein subunits (ribosomes, spliceosomes etc.). It has even been proposed that all proteins in a given cell are connected in a huge network in which certain protein interactions are forming and dissociating constantly.

In Figure 1, complex networks showing the interactions among proteins help scientists understand how a drug affecting one protein will affect overall cell functioning. This protein network for Brewer's yeast shows which proteins are critical for survival (red), which are important for growth but not critical to survival (orange), which can be removed without slowing growth or killing the cells (green), and which are of unknown importance (yellow).

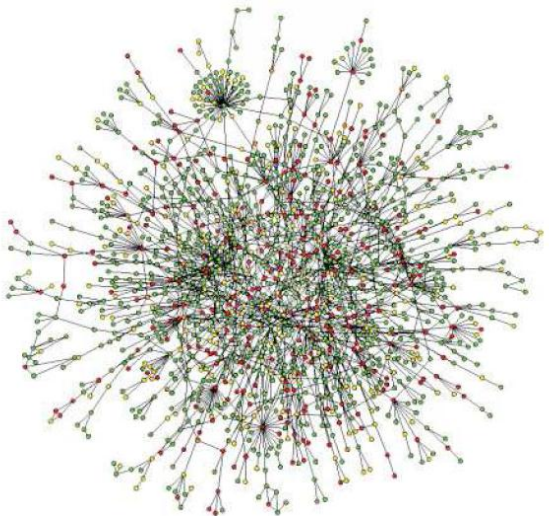

Figure 1: Protein network for Brewer's yeast.

Disease-Causing Genes –

The information contained in our genes is so critical that simple changes can lead to a severe inherited disease, make us more inclined to develop a chronic disease, or make us more vulnerable to an infectious disease. In Figure 2, the rules of governing monogenetic diseases and complex diseases are showed. While monogenetic diseases are caused by a single gene and complex diseases are complicated combinations of many genes.

Scientists currently believe that single gene mutations cause approximately 6,000 inherited diseases. These diseases are called single gene or monogenic diseases because a change in only one gene causes the disease. These diseases include a number of lung and blood disorders, such as cystic fibrosis, sickle cell anemia, and hemophilia. Although these conditions are not popular however they still affect millions of people worldwide. The rules that underlie the inheritance of major common diseases are not as straightforward. These diseases include heart disease, diabetes, Alzheimer disease, psychiatric disorders, and osteoarthritis.

These common diseases result not just from a change in one or a few genes, but from a combination of the effects of the environment and a number of susceptibility genes.

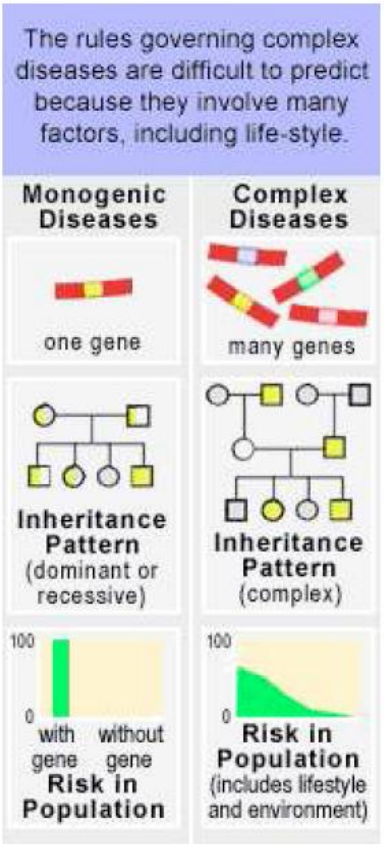

Figure 2: Monogenetic diseases and complex diseases.

Susceptibility genes contribute to an individual's risk of developing a specific disease, but usually are not enough to cause the disease. Susceptibility genes may influence the age of onset of a disease, contribute to its rate of progression, or help to protect against it. Understanding the rules of their inheritance and their roles in disease is not a simple task. Different alleles may be associated with different degrees of susceptibility, or risk. The APOE gene on chromosome 19 is one example of a disease susceptibility gene.

An individual who has two copies of one variant allele of APOE is more likely to develop Alzheimer disease at an earlier age than an individual with a different APOE genotype.

From Protein-Protein Interaction Networks To Disease-Causing Genes -

Interactions between specific pairs or groups of proteins are essential to all stages of development and homeostasis. Not surprisingly, many human diseases can be traced to aberrant protein?protein interactions, either through the loss of an essential interaction or through the formation of a protein complex at an inappropriate time or location. There are some relationships between diseases and protein interactions such as a hitchhiker's guide to pathogen host interactions, normal protein-protein interactions gone wrong, protein protein interaction inhibitors.

THE ROLE OF NETWORKS IN MEDICINE

Networks provide a systems-level understanding of themechanisms underlying diseases by serving as a model for data integration and analysis.

They have been used to gain insight into disease mechanisms, study comorbidities, analyze therapeutic drugs and their targets and discover novel network-based biomarkers. Network science deals with complexity by "simplifying" complex systems to components (nodes) and interactions (edges) between them (Fig. 3). These simplifications help researchers make useful discoveries.

Networks can be constructed purely based on gene expression information, including transcriptional regulatory networks and co-expression networks, or can also be built upon prior knowledge of protein–protein interactions. The nodes in a network representation are metabolites or macromolecules such as proteins, RNA molecules and gene sequences, while the edges are physical, biochemical or functional interactions. The resulting "interactome" network can serve as scaffold information to extract global or local graph theory properties,which lead to a better understanding of biological processes.

Since cellular networks consist of various types of interaction and regulation, networks reflecting this complex scenario will provide better insight into the problem in hand.

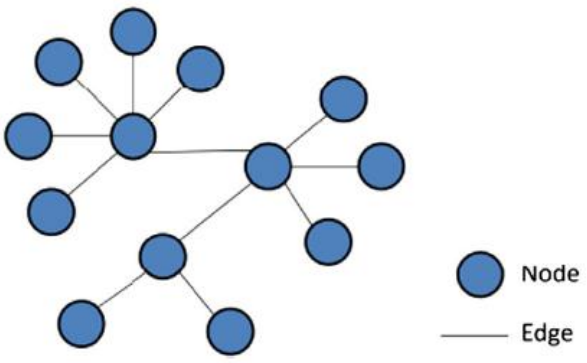

Fig. 3. Simple representation of a network, nodes representing components and edges representing interactions.

Regulatory interaction networks, metabolic networks, signaling networks, and protein–protein interaction networks cannot be considered in isolation or as independent entities. Rather, we have to incorporate their intricate interwoven structure. Proteins might act alone or in combination: as transcription factors and regulators of protein abundances, as enzymes they

catalyze and coordinate the basic cellular metabolic processes, and they react to external and internal stimuli activating other proteins in signaling cascades.

All of these processes in turn provide cues that may lead to the formation or termination of protein interactions and complexes. PPI networks in particular have become a valuable resource in this context.

PPI AND DISEASE EXAMPLE

Let us consider cerebral malaria as an example to understand how an analysis of PPI could be used to elucidate the molecular basis of disease. Here, a wide range of experimental and predicted human–Plasmodium (host-parasite), human-human (host-host) and Plasmodium-Plasmodium (parasite-parasite) PPI are combined and analyzed in the context of key events and processes of cerebral malaria, a dangerous infectious disease (Rao et al., 2010).

Cerebral malaria is a severe form of malarial infection, characterized by cerebral complications, such as neuronal damage and coma (Moxon et al., 2009). The disease is characterized by processes such as sequestration of infected red blood cells to cerebral capillaries and venules, systemic inflammation, hemostasis dysfunction and neuronal damage (Wilson et al., 2008).

An automated literature retrieval module was developed using Entrez Programming Utilities (Sayers et al., 2010) to retrieve the list of full-text articles relevant to the malarial parasite. This article set was pruned using the Medical Subject Headings (MeSH) controlled vocabulary for articles relevant to cerebral malaria. The resultant set was augmented by articles retrieved from the Google Scholar database using appropriate disease-specific query terms such as systemic inflammation, hemostasis dysfunction etc. This article corpus had two main uses:

- For extracting biochemical and signaling events of relevance in cerebral malaria.
- Identifying pairs of interacting proteins within the host, within the parasite and between host and parasite.

Gene Ontology (GO) cellular component annotations from PlasmoDB, a comprehensive Plasmodium resource, were used to prune the unified PPI dataset using the approach of Mahdavi & Lin (2007). In the case of PPI involving parasite proteins, only those proteins that were annotated to be present on the parasite surface or were reported to be released during the relevant stage of the parasite were considered. For the human protein annotations, tissue-specific annotations from UniProt (Hubbard et al., 2009) were used in the pruning process.

The resultant PPI subset was then analyzed by mapping the PPI to key events that influence the processes of the disease, as identified from the key review articles. The analysis showed the potential significance of apolipoproteins and heat-shock proteins on efficient Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) presentation, role of the merozoite surface protein (MSP-1) in platelet activation, the role of albumin in astrocyte dysfunction and the effect of parasite proteins in transforming growth factor (TGF)-β regulation. The linking of these PPI to molecular events associated with the disease pathogenesis provides a basis for further experiments to determine the molecular basis of this fatal disease.

METHODOLOGY

In order to identify associations between diseases by mapping their respective protein interaction networks with statistical significance values, we took the following steps.

Extraction of human protein-disease relationships was achieved though Structured Query Language querying of the PhenoGO database. We extracted all UMLS-coded diseases classified under the “Disease” semantic type hierarchy along with their associated proteins. In this study, we chose to stay on a more conservative side, and only extracted diseases associated with more than 4 proteins to avoid errors stemming from mis-assignment in PhenoGO and to reduce spurious predictions in the next step from the hypergeometric distribution because a single error contributes proportionally to a larger statistical impact on a smaller sample of protein in the statistical method that follows (equation 1). These UMLS-coded terms fall under the UMLS semantic types ‘Congenital Abnormality’, ‘Disease or Syndrome’, ‘Experimental Model of Disease’, ‘Anatomical Abnormality’, and ‘Neoplastic Process’. The resultant set consists of 154 diseases and their 1,931 associated proteins (http://phenos.bsd.uchicago.edu/PSB2007/).

Integration and Discovery. The second step is to correlate diseases with their underlying protein-protein interaction networks using a statistical approach. In this study, we used the Reactome protein interaction dataset to define the underlying topological networks associated with these diseases. The common proteins between diseaseassociated proteins in PhenoGO and proteins in the Reactome were identified by using the identifiers in the UniProt. The Reactome data set defines four distinct types of reactions: 1) neighboring reactions, which define interactions that occur consecutively; 2) indirect complexes, which define interactions which involve

subcomplex interaction, but not direct binding/interaction; 3) direct complex, defining protein-protein complexes; and 4) reaction, representing situations where the two proteins participate in the same reaction. The Reactome dataset was normalized to a set of paired Swiss-Prot accession numbers, and filtered to remove neighboring reactions and indirect complexes, leaving only entries for binary interactions and direct complexes. This data set contains 20,317 distinct interactions corresponding to 1,140 distinct proteins. From the 154 diseases, we generated combinations of pairs of diseases, and for each pair of diseases, proteins in both diseases were also paired for all potential combinations. These protein pairs were then cross-referenced with our filtered Reactome data set to determine if they participated in reactions or formed direct complexes with one another. There are two basic types of relationships used in calculations in our methods. These relationships correspond to the two scenarios we considered to determine whether two diseases share interaction networks: 1) an identity relationship where common proteins are shared by two diseases, and 2) direct interactions between protein A in one disease and protein B in the other disease. As related diseases can share both types of relations, and due to the requirements of the hypergeometric distribution, we consider both in the underlying protein-protein interaction network in diseases. Based on this, we calculated the correlations between all possible pairs of diseases by applying the hypergeometric distribution function to identify significantly correlated diseases (equation 1) and adjustments for multiple a posteriori comparisons (equation 2), as shown below:

$$p(i >= m \mid N, M, n, m) = \sum_{i=m}^{n} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}} \quad (1)$$

In equation 1, 'N' represents the total number of all pair combinations between proteins of any two diseases in the experiment that includes the possibility of sharing the same proteins (identical protein pair between two diseases), 'M', as the sum of number of observed distinct pairs of interacting proteins that exist in the Reactome database for all

the diseases in the experiment (direct interaction only), 'n' as the putative total number of pairs of proteins that could exist in a pair of disease, and 'm" as the sum of the observed number of common proteins shared between two specific diseases and the number of distinct pairs of interacting proteins observed in the Reactome database for these two specific diseases (M ∩ n). This measure gives a p-value which is then adjusted for multiple comparisons with the Dunn-Sidak method (a derivative of the Bonferroni method):

$$p' = 1 - (1 - p)^r \quad (2)$$

In equation 2, p' and p represent the corrected and uncorrected p-values, respectively, and r represents the number of independent comparisons, which is the number of disease pairs (r=11,703) used in the study. These corrected p-values are then thresholded at $p<0.05$ to determine the final set of significantly correlated disease-disease relationships. Multiple diseases and genes sharing the same PubMed IDs can artificially boost the statistical significance of these disease pairs, therefore relationships mapping to more than 2 overlapping PubMed IDs were removed to reduce the this artifact. A total of 11,703 disease pairs passed the filter out of 11,780 candidates. 77 combinations had more than two PMID overlaps and were filtered out as a result of this process.

Evaluations. Two evaluations were conducted. The first one, a quantitative evaluation, was designed to control for the error rate in either assigning a protein disease relationship in PhenoGO or a protein-protein interaction in Reactome. It consisted of establishing the reliability of the predictions if we introduced noise in the integrated database network (10% more protein-protein interactions in the same set of diseases).

The second one, a qualitative evaluation, consisted of carefully examining the discovered protein interactions shared by two diseases and identifying references in the scientific literature that validate the phenotypic overlap and potentially the protein interactions.

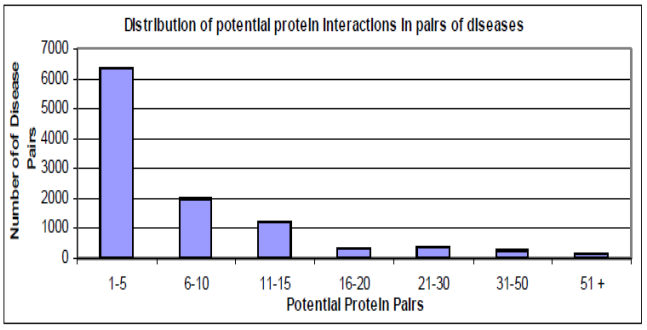

Figure 4. Distribution of the number of disease pairs from PhenoGO according to the number of possible protein interactions observed between their proteins in the Reactome.

RESULTS

In this study, we examined a subset of PhenoGO pertaining to human diseases in order to identify relationships between these diseases according to criteria described in the methods. This filtering resulted in a set of 154 diseases and their 1,931

associated proteins. The intersection between the proteins of the Reactome and those of PhenoGO further reduced the set of proteins to 286. The number of candidate proteins per disease was greatly reduced by the need to be present in the Reactome dataset, and therefore the totals are smaller than observed in the PhenoGO database alone. We lose approximately

70% of the proteins in this process due to the limited content of the Reactome. In order to identify relationships between these diseases, we analyzed their underlying proteinprotein interaction maps by applying a statistical method (details of equations in the Method Section). Of the 154 selected diseases, there are (285*286/2+286) = 41,041 distinct combinations of protein pairs and identical protein overlap (term N, equation 1) possible for all possible disease pairs, of which only 4,857 exist in the Reactome (term M, equation 1). Figure 4 summarizes the distribution of protein-protein pairs per combination of diseases in our set. In ~60% of the 11,703 disease pairs under consideration, the number of potential protein-protein interactions is five or less (no significant predictions from this category), and about 40% of them have more than five interactions. We then proceeded in calculating the correlation between groups of pairs of interacting proteins associated with every pair of diseases according to equations 1 and 2 (file available at http://phenos.bsd.uchicago.edu/PSB2007/). Based on the correlations of the shared protein interacting pairs between diseases, we identified 10 pairs of diseases that are significantly correlated due to their shared proteins and protein-protein interactions out of 11,703 disease pairs examined in this study (Table 1).

UMLS ID	Disease 1	UMLS ID	Disease 2	P-PI (#)	pvalue	Corrected pvalue	Ref
C0009207	Cockayne Syndrome	C0043346	Xeroderma Pigmentosum	38	7.3e-22	8.5e-18	[31]
C0043346	Xeroderma Pigmentosum	C0085390	Li-Fraumeni Syndrome	24	6.7e-11	4.9e-06	[32]
C0007001	Carbohydrate Metabolism, Inborn Errors	C0002514	Amino Acid Metabolism, Inborn Errors	9	8.3e-10	6.2e-05	*
C0009404	Colorectal Neoplasms	C0950123	Genetic Diseases, Inborn	16	6.7e-10	5.0e-05	[33]
C0085390	Li-Fraumeni Syndrome	C0009207	Cockayne Syndrome	16	2.7e-09	1.9e-04	NA
C0009404	Colorectal Neoplasms	C0015625	Fanconi's Anemia	8	1.5e-05	6.7e-01	[9]
C0009404	Colorectal Neoplasms	C0085413	Polycystic Kidney, Autosomal Dominant	8	1.5e-05	6.7e-01	[21]
C0024141	Lupus Erythematosus, Systemic (LES)	C0004364	Autoimmune Diseases	4	9.3e-05	9.9e-01	*
C0024314	Lymphoproliferative Disorders (LD)	C0004364	Autoimmune Diseases	6	1.3e-04	9.9e-01	[34]
C0024314	LD	C0024141	LES	6	1.3e-04	9.9e-01	[35]

* self-evident relations between disease pairs. Ref = references confirming the predictions, P-PI = Protein –Protein Interaction #

Table 1. Top ranked significantly correlated diseases.

DISCUSSION

The protein-protein interaction network constructed by the Reactome dataset provides us a framework for structuring the knowledge of human diseases, which enables an objective approach to examine the molecular underpinnings of diseases in the context of their known molecular interactions on genomic scale. This method not only allows us to conduct high throughput computational analysis of the relations between diseases, but also reveals the underlying molecular relationships between diseases. Furthermore, new relationships between well-known diseases and new diseases could be revealed based on their overlapping molecular networks.

CONCLUSION

We developed and evaluated an automatic system to predict protein interactions shared by two or more diseases. It augments current protein interaction networks by integrating literature-based knowledge of protein-disease associations and systematically identifying the statistically significant Protein Interactions of Diseases (PID). Results demonstrated that the PID system provides accurate predictions and is scalable in a number of dimensions: (i) it enables high throughput predictions, and (ii) it scales across different protein-interaction datasets. Beyond direct protein-protein interactions, it also provides the theoretical framework to compare shared pathways between diseases. In the future, this framework could be applied to more complex diseases to determine if their shared phenotypes are a result of the shared molecular mechanism and pathways.

We developed and evaluated an automatic system to predict protein interactions shared by two or more diseases. It augments current protein interaction networks by integrating literature-based knowledge of protein-disease associations and systematically identifying the statistically significant Protein Interactions of Diseases (PID). Results demonstrated that the PID system provides accurate predictions and is scalable in a number of dimensions: (i) it enables high throughput predictions, and (ii) it scales across different protein-interaction datasets. Beyond direct protein-protein interactions, it also provides the theoretical framework to compare shared pathways between diseases. In the future, this framework could be applied to more complex diseases to determine if their shared phenotypes are a result of the shared molecular mechanism and pathways.

REFERENCES

Benjamin Schuster Bockler and Alex Bateman (2008). Protein interactions in human genetic diseases. Genome Biology, 9(1):R9.1–R9.12.

Berg, J., M. LassigA. Wagner, Structure and evolution of protein interaction networks: a statistical model for link dynamics and gene duplications. BMC Evol Biol, 2004. 4(1): p. 51.

Bertola, D.R., C.A. Kim, L.M. Albano, H. Scheffer, R. Meijer H. van Bokhoven, Molecular evidence

that AEC syndrome and Rapp-Hodgkin syndrome are variable expression of a single genetic disorder. Clin Genet, 2004. 66(1): p. 79-80.

Hubbard TJ, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Graf S, Haider S, Hammond M, Holland R, Howe K, Jenkinson A, Johnson N, Kahari A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Lawson D, Longden I, Megy K, Meidl P, Overduin B, Parker A, Pritchard B, Rios D, Schuster M, Slater G, Smedley D, Spooner W, Spudich G, Trevanion S, Vilella A, Vogel J, White S, Wilder S, Zadissa A, Birney E, Cunningham F, Curwen V, Durbin R, Fernandez-Suarez XM, Herrero J, Kasprzyk A, Proctor G, Smith J, Searle S & Flicek P. (2009). Ensembl 2009. Nucleic Acids Research 37: D690-697.

J.Y. Chen, C. Shen, and A.Y Sivachenko (2006). Mining Alzheimer Disease Relevant Proteins from Integrated Protein Interactome Data. In Pacific Symposium on Biocomputing, volume 11, pp. 367–378.

Jansen R, Yu H, Greenbaum D, Kluger Y, Krogan NJ, Chung S, et al. (2003). A Bayesian networks approach for predicting protein-protein interactions from genomic data. Science. 302(5644): pp. 449–453. [PubMed: 14564010]

Jansen, R.H. YuD. GreenbaumY. KlugerN.J. Krogan S. Chung, et al. (2003). A Bayesian networks approach for predicting protein-protein interactions from genomic data. Science, 302(5644): p. 449-53.

Lubovac, Z., J. GamalielssonB. Olsson (2006). Combining functional and topological properties to identify core modules in protein interaction networks. Proteins, 2006.

Mahdavi MA & Lin YH. (2007). False positive reduction in protein protein interaction predictions using gene ontology annotations. BMC Bioinformatics 8: p. 262.

Moxon CA, Heyderman RS & Wassmer SC. (2009). Dysregulation of coagulation in cerebral malaria. Molecular and Biochemical Parasitology 166: pp. 99-108.

Nabieva, E., K. Jim, A. Agarwal, B. Chazelle M. Singh (2005). Whole-proteome prediction of protein function via graph-theoretic analysis of interaction maps. Bioinformatics, 21 Suppl 1: pp. 302-10.

Rzhetsky A, Gomez SM. (2001). Birth of scale-free molecular networks and the number of distinct DNA and protein domains per genome. Bioinformatics. 17(10): pp. 988–996. [PubMed: 11673244]

Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Lu Z, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Mizrachi I, Ostell J, Panchenko A, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Slotta D, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Wang Y, John Wilbur W, Yaschenko E & Ye J. (2010). Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 38:D5-16.

Wilson NO, Huang MB, Anderson W, Bond V, Powell M, Thompson WE, Armah HB, Adjei AA, Gyasi R, Tettey Y & Stiles JK. (2008). Soluble factors from Plasmodium falciparum-infected erythrocytes induce apoptosis in human brain vascular endothelial and neuroglia cells. Molecular and Biochemical Parasitology 162: pp. 172–176.

Zenteno JC, Venegas C, Kofman-Alfaro S. (1999). Evidence that AEC syndrome and Bowen—Armstrong syndrome are variable expressions of the same disease. Pediatr Dermatol. 16(2): pp. 103–107. [PubMed: 10337671]

Corresponding Author

Rajeev Kumar*

Research Scholar, SSSUTMS, Sehore

E-Mail – chairman.iab@gmail.com