Inferring Human Phenotype Networks from Genome-Wide Genetic Associations

Christian Darabos[1,⋆], Kinjal Desai[1,*], Richard Cowper-Sal·lari[1], Mario Giacobini[2], Britney E. Graham[1], Mathieu Lupien[3], and Jason H. Moore[1]

[1] Department of Genetics, The Geisel Medical School at Dartmouth College, Lebanon, NH 03756, U.S.A.
[2] Computational Epidemiology Group, Department of Veterinary Sciences, and Complex Systems Unit, Molecular Biotechnology Center, University of Torino, Italy
[3] Ontario Cancer Institute, Princess Margaret Cancer Center-University Health Network, Ontario Institute for Cancer Research and the Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

Abstract. Networks are commonly used to represent and analyze large and complex systems of interacting elements. We build a human phenotype network (HPN) of over 600 physical attributes, diseases, and behavioral traits; based on more than 6,000 genetic variants (SNPs) from Genome-Wide Association Studies data. Using phenotype-to-SNP associations, and HapMap project data, we link traits based on the common patterns of human genetic variations, expanding previous studies from a gene-centric approach to that of shared risk-variants. The resulting network has a heavily right-skewed degree distribution, placing it in the *scale-free* region of the network topologies spectrum. Additional network metrics hint that the HPN shares properties with *social networks*. Using a standard community detection algorithm, we construct *phenotype modules* of similar traits without applying expert biological knowledge. These modules can be assimilated to the disease classes. However, we are able to classify phenotypes according to shared biology, and not arbitrary disease classes. We present a collection of documented clinical connections supported by the network. Furthermore, we highlight phenotypes modules and links that may underlie yet undiscovered genetic interactions. Despite its simplicity and current limitations the HPN shows tremendous potential to become a useful tool both in the unveiling of the diseases' common biology, and in the elaboration of diagnosis and treatments.

1 Introduction

Biology at the system's level is a holistic approach to the study of an organism's entire phenotypes. When applied to humans, systems biology encompasses all aspects, both environmental and internal, of an individual to understand its traits and diseases. It offers the promise of personalized diagnostics, prognostics and medical treatments [14]. Because of the sheer complexity and the number of interactions, the preferred visualization method of systems biology is the network. Indeed, networks offer relatively straight

⋆ First Co-authors.

L. Vanneschi, W.S. Bush, and M. Giacobini (Eds.): EvoBIO 2013, LNCS 7833, pp. 23–34, 2013.

forward and intuitive representations of interaction phenomena, and allow sophisticated statistical analysis of their intrinsic properties.

In this work, we focus on the system-wide relationships between human phenotypic traits (PT), encompassing physical attributes (e.g. eye color, waist circumference), diseases (e.g. coronary heart disease, Type 1 and 2 diabetes), and behavioral characteristics (e.g. smoking behavior). Elucidating relationships between human traits or diseases is becoming increasingly important in the study of complex genetic disorders. These traits are related through shared genes, proteins and possibly regulatory elements. Identifying these links may help reveal shared mechanisms driving the set of connected diseases. Ultimately, a thorough understanding of these connections may provide the clinical tools necessary to design common drug targets. The potential biological and clinical outcomes justifies efforts to study the phenotype genotypical interactions. These interactions are mathematically and visually represented as a graph: the Human Phenotype Network (HPN). Previous network-based studies of diseases have proven useful for visualizing large disease datasets grouped by common mutated genes, similar gene expression profiles or shared protein interactions [7,24,3]. However, a gene-centric focus has biased the generation and interpretation of these networks, given that coding regions constitute less than 2% of the entire human genome. Genome Wide Association Studies (GWAS) have identified genetic predispositions to disease using a non-candidate-driven approach. To date, approximately 6,000 single nucleotide polymorphisms (SNPs) have been reported as genetic risk-variants for about 600 diseases and traits. Over 90% of risk-associated SNPs (raSNPs) identified by the GWAS fall outside of coding regions ([8]), stressing the requirement for a more global assessment of shared risk-variants. Here we propose a non-gene centric method, relying on genetic risk factors, such as SNPS, and construct a network of traits and diseases based on their shared GWAS loci. Previous studies also focused on classifying diseases into arbitrary disease classes usually based on the organs or the physical location of the disease in the human body, disregarding the shared biology of the diseases. Here, we take a different approach, classifying phenotypes into modules, by using a community detection algorithm based on the phenotypes' position with respect to one another within the HPN. We present a collection of clinical interactions that are corroborated by the network (Fig. 1) and we show that the HPN reveals phenotypes sharing loci that may underlie as yet uncharacterized interactions.

2 Background

In this section, we define the fundamental concepts used in the methods section to build the HPN (Section 3);

2.1 Genome-Wide Association Studies

Genome-wide association studies (GWAS) identify common genetic variants, such as single-nucleotide polymorphisms (SNP), found in the genotype of different individuals in association with phenotypical traits. A SNP is said to be associated with a trait if it is more prevalent in the group presenting the phenotype of interest (cases), when compared to the group not presenting it (controls). SNPs associated with a trait, or risk-associated SNPs (raSNPs), mark the region of the human genome that is believed to

influence the probability (or risk) of the trait's occurrence in an individual [15]. Pairs or groups of SNPs are said to be in *linkage disequilibrium* when they are found to occur together more (or less) often than would be expected at random [6]. The catalog of published GWAS maintained by the National Human Genome Research Institute (NHGRI) at the National Institute of Health (`http://www.genome.gov/gwastudies/`) aggregates studies that report phenotype-to-raSNP(s) associations. The NHGR catalog used in this study, dated 05/17/11, and our primary source of raSNP-trait association data, reports over 600 PTs associated with approximately 6,000 raSNPs.

Imputed Risk Associated Variome. For each trait in the catalog, we extract the complete set of raSNPs, which we call a risk-associated variome (RAV). To address the low genomic coverage provided by GWAS, we associate each raSNP with all SNPs found in linkage disequilibrium (ldSNPs) [6] using the HapMap project data [9]. SNPs in linkage disequilibrium form clusters of variants that statistically appear in the same patient. The HapMap project aims at building a repository of describing the common patterns found in human genetic variations (`http://hapmap.ncbi.nlm.nih.gov/`). The resulting imputed variome (iRAV) will allow us to establish connections between diseases/traits that share blocks, i.e. that have overlapping iRAVs. A recent study [25] shows that SNPs in linkage disequilibrium (ldSNPs) with prostate cancer risk-associated SNPs modulate the expression of an oncogene by altering transcription factor binding sites. The inclusion of ldSNPs in our analysis is therefore expected to be valuable.

2.2 Networks

As previously mentioned, network theory can provide powerful tools for visualizing complex systems. Networks are being used with increasing frequency to analyze large scale systems, such as the Human Disease Network (HDN), which will be the focus of this study. A network can take an extraordinarily complex system and reduce it to a relatively simple form, revealing underlying connections and important clustering details that ordinarily would not be seen, when studying individual or non-complex relationships between traits [16]. Intuitively, a network is a collection of nodes and the edges connecting them. The degree of a node is determined by the number of edges that are attached to it [16]. The degree distribution of a network defines the probability that each node will have a certain degree. The plot of the degree distribution probability function informs us of important global properties of the network. For instance, if the plot curve follows a normal or a Poisson distribution, then the network's topology is said to be *random*. On the other hand, if the plot is right-skewed with a long tail, this indicates that most of the nodes in the network are of a low degree with a few highly connected nodes that are referred to as hubs. This type of network is called *scale-free* and its degree distribution tends to follow a power-law, or decaying exponential curve. Most biological networks are found to be in the scale-free family. When the degree distribution of a scale-free network is plotted on a logarithmic scale, the resulting curve is approximately linear across the top [16]. In the case of relatively small networks, it is impossible to affirm the presence of a scale-free network. We can, at best, show the existence of a power-law type degree distribution, and not dismiss the scale-free hypothesis.

Modules within a Nework. The clustering coefficient (CC) of a network measures the degree to which nodes tend to form closely knit communities with a higher than average connectivity [23]. The CC of networks found in nature, in particular social and biological networks, show a higher degree of clustering than that observed in randomized networks of identical size. This measurement allows one to identify clusters of nodes within the network that are likely to share common attributes based on the structural properties of the network, without using any specific information about the nature of the nodes themselves. These clusters are called modules in the case of general networks, and communities in social networks. The Louvain method of community detection in large scale networks, based on a greedy optimization method [5], is a widely accepted algorithm to build communities (or modules) within a network with no expert-knowledge.

2.3 Human Disease Networks

In recent years there has been a trend toward studying disease through network based analysis of various systems of connections between diseases. The result is the Human Disease Network (HDN). The nodes in the HDN represent human genetic disorders and the edges represent various connections between disorders, such as gene-gene or protein-protein interactions, to name only a few. The HDN is helpful in visualizing human disorders and their corresponding interactions on a large scale, which gives us the opportunity to see the relationships between disorders. The underlying connections of the HDN contribute to the understanding of the basis of disorders, which in turn leads to a better understanding of human diseases.

One study by Goh, *et al.*[7], explored the HDN built on mutated genes shared by different diseases. Another study, which is similar in some ways to ours, by Li *et al.*[13] traced the raSNPs connecting disease traits. In 2009, Silpa Suthram *et al.*[21] found that when diseases were compared and contrasted by an analysis of disease-related messenger RNA (mRNA) expression data and the human protein interaction network, there were significant similarities between certain diseases and that some of the correlated diseases shared drug treatments, as well. This could help us target certain genes for treatment. In 2009, Barrenas *et al.*[3] further studied genetic architecture of complex diseases, by doing a GWAS, and found that complex disease genes are less central than the essential and monogenic disease genes in the human interactome. In the present work, we expand our study to include not only disease traits, but also behaviors and normal variations in humans, such as hair color, and explore large portions of non-coding variations in the human genome. In addition, we include not only raSNPs, but also ldSNPs to achieve a better coverage of the phenotype interactions.

3 HPN Based on Genetic Variations

This section describes our proposed method to construct a HPN of PTs and diseases based on their shared GWAS loci. This model includes data from hundreds of GWAS studies, all catalogued by the NHGRI, and adds the HapMap project data to build comprehensive clusters of rare variants for each phenotype (iRAVs). As we will show in

Section 4, this approach offers interesting insight into the way phenotypes may be linked by common genetic variations. The network is built following the steps below; each step of this algorithm is represented in Fig. 1 below.

1. from the NHGR catalog, extract all PTs associated with at least one raSNP in at least one study, and set those as nodes;
2. associate each PT with the RAV containing all the raSNPs identified;
3. extend each RAV to its iRAV by building clusters of ldSNPs around each raSNP;
4. identify overlapping iRAVs, and connect the associated PTs in the HPN with a directed edge;
5. set the edge weight as the normalized number of iRAVs shared by the 2 PTs, i.e. the number of overlapping iRAVs over the total number of iRAVs associated with the source vertex of that edge;

As a result of Step (1), the network will not contain any isolated nodes. We are only interested in PTs that have been associated with raSNPs, and their possible shared biology. The original NHGR database contains 646 PTs; by removing the isolate nodes, the HPN contains 401 nodes connected to at least 1 other node.

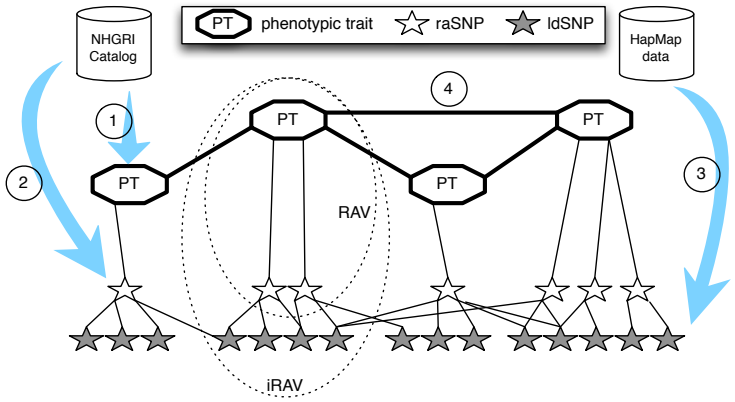

Fig. 1. Step-by-step description of the method to obtain the HPN. The circled numbers correspond to the steps of the method described above.

The resulting network is shown in Fig. 2, where the nodes represent the PTs, and the edges correspond to overlapping iRAVs. To increase the readability, nodes and edges were filtered (see legend of Fig. 2). All the statistics below are, however, computed on the complete (unfiltered) network.

In Fig. 2, the nodes and labels sizes are proportional to the original degree of the PT (before filtering). The edge width is in turn proportional to the number of SNP clusters overlapping in the nodes' iRAVs (weight). Visually, we notice the network is composed of a small number of highly connected hubs: coronary heart disease, cholesterol, Crohn's disease. The vast majority of the nodes are, however, very sparsely connected,

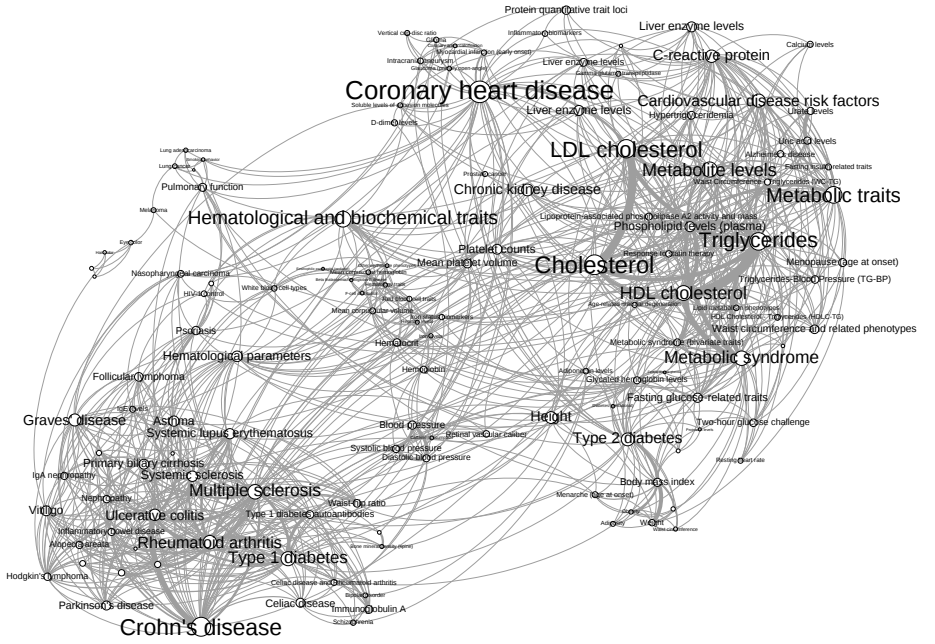

Fig. 2. Human Phenotype Network. In order to increase the readability, we have filtered out nodes with a degree smaller than 5 (i.e. connected to less than 5 other nodes), showing only 137 nodes (about 30%), and edges with a weight lower than 2 (i.e. connecting PTs that have iRAVs overlapping by less than 2 ra/ldSNP clusters), showing about 45% of the actual edges. To further facilitate the readability, we have manually merged a number of clearly redundant nodes and depicted the double directed edges as single undirected.

allowing us to speculate on the scale-free nature of the network. Our hypothesis is supported by the degree distribution plots in Fig. 3. Indeed, the degree distribution is clearly right-skewed, with a *heavy-tail*.

Scale-free networks are ubiquitous in nature and in biology[2], and our HPN is no exception. Crohn's disease and Coronary Heart disease are the main hubs of the network with connections to over 60 other traits. They are followed by Hematological and biochemical traits and LDL Cholesterol related phenotypes. Table 1 summarizes a number of the standard network properties and statistics computed on the HPN. For comparison purposes, we have also included the statistics of the HPN when we disregard ldSNPs, using direct raSNP overlap only. The results are clearly in favor of including ldSNPs into our study, as it offers a more complete view of possible phenotypic interactions.

Indeed, the complete HPN includes about 100 more traits, and about 3 times the number of edges. The HPN's clustering coefficient of 0.595 is much higher than that expected of random networks (RN) of identical dimensions ($CC = 0.035$). The short average path length implies a large number of shortcuts across the network[22]. Moreover,

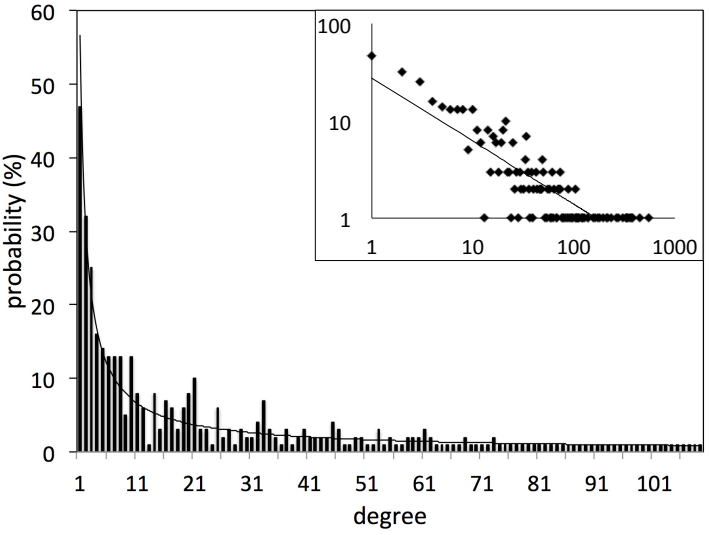

Fig. 3. Degree distribution of the HPN. The vertical axis represents the probability (in percents) of a node having the corresponding degree on the horizontal axis. The inset figure is plotted on a logarithmic scale. The trend lines are shown to offer an approximation of the *long-tailed* function of the distribution and show that the distribution is close to the inverted power-law function of a scale-free distribution.

Table 1. Properties and Statistics of the HPN

Property / Statistic	complete HPN	raSNPs-only HPN
#nodes	401	295
#edges	2845	932
#components	9	25
largest component	385 nodes, 2837 edges	252 nodes, 989 edges
average degree	14.19	6.39
average weighted degree	37.54	10.03
average clustering coefficient	0.595	0.427
average path length	2.961	3.70

the largest connected component (LCC) is significantly smaller than the HPN, where the LCC of a RN with a similar average degree would englobe all nodes. Together, these three properties place the HPN in the *social network* range, where clusters of individuals tend to form with a higher-than-random probability. This also hints that interesting insights can be gained by using a clustering algorithm to identify the HPN's intrinsic communities. In Section 4, we discuss the results of the clustering algorithm (see 2.2), and study the biological and clinical implications that can be gathered from the HPN.

4 Biological and Clinical Implications

In this section, we analyze the constructed HPN and present a collection of clinical interactions that are corroborated by the network (Fig. 4). Furthermore, the network reveals phenotypes sharing loci that may underlie yet uncharacterized interactions. The inclusion of behaviors and traits, as opposed to diseases only, will prove very informative. We use the Louvain method (see Section 2.2) to build the network statistics based communities within our HPN [5]. Fig. 4 offers an overview of the network with a color-coding of the 24 different modules identified by the algorithm. In [7], the authors have manually classified the diseases present in the HDN into "disease classes", leaving an important number "unclassified". Our modules based approach classifies all genetically related phenotypes automatically into approximate classes, based on their linkage within the HPN. This contrasts strongly with previous work, where arbitrary disease classes grouped phenotype regardless of their shared biological attributes. For example, in previous studies, all cancers were part of the same class, regardless of the cancer type. Using the phenotype modules and the community detection algorithm, our framework is able to classifies each cancer with the phenotype that they are most likely to share genetic attributes with. Because of the large differences in the data sets, notably the addition of physical and behavioral traits, we cannot directly compare the classes used in the HDN and our HPN modules. The largest clusters appear around the hubs: "Crohn's disease" and "Type 1 diabetes" in Fig. 4 cluster A; "Coronary heart disease" and "Hematological and biochemical traits", Fig. 4 cluster B; "C-reactive protein" and "Chronic kidney disease", Fig. 4 cluster C; "LDL cholesterol" and "Triglycerides" (metabolic diseases); and "Type 2 diabetes" and "Obesity" (metabolic).

The HPN presents several edges that confirm well-characterized genetic interactions. These include the dense interconnectivity between immune-related disorders and phenotypes. For instance, systemic sclerosis and rheumatoid arthritis, both autoimmune disorders in the systemic inflammatory rheumatic disease family, are connected and are part of the same module (Fig. 5B). The metabolic diseases centered around excessive weight, elevated body mass index, obesity and Type 2 diabetes also form a module (Fig. 5C).

The HDN also points to connections between diseases known to rely on common factors. For instance, bone mineral density (hip and spine) is linked with inflammatory diseases such as Crohns disease and Ulcerative Colitis, both subsets of Inflammatory Bowel Disorder (IBD). Nuclear factor kappa B (NFκB) is known to be involved in driving IBD [10] and has recently been shown to play a role in regulating genes responsible for bone formation [12] a key factor in establishing bone mineral density. Bone mineral density (hip and spine) is also connected with breast cancer (not shown here). It is well established that the estrogen receptor alpha (ESR1) drives oncogenesis in over two-thirds of all breast cancers [17]. Mutations in the estrogen receptor gene have also been associated with loss of bone mineral density in humans [20]. The HDN also reveals connections between behavioral traits and diseases. For instance, lung cancer is connected with smoking behavior and nicotine dependence within the same module (Fig. 5B). This raises the possibility that a SNP associated with a certain disease phenotype may not affect the biology in the disease tissue but instead promote behaviors that increase the risk of the connected disease. The HDN also shows a connection between

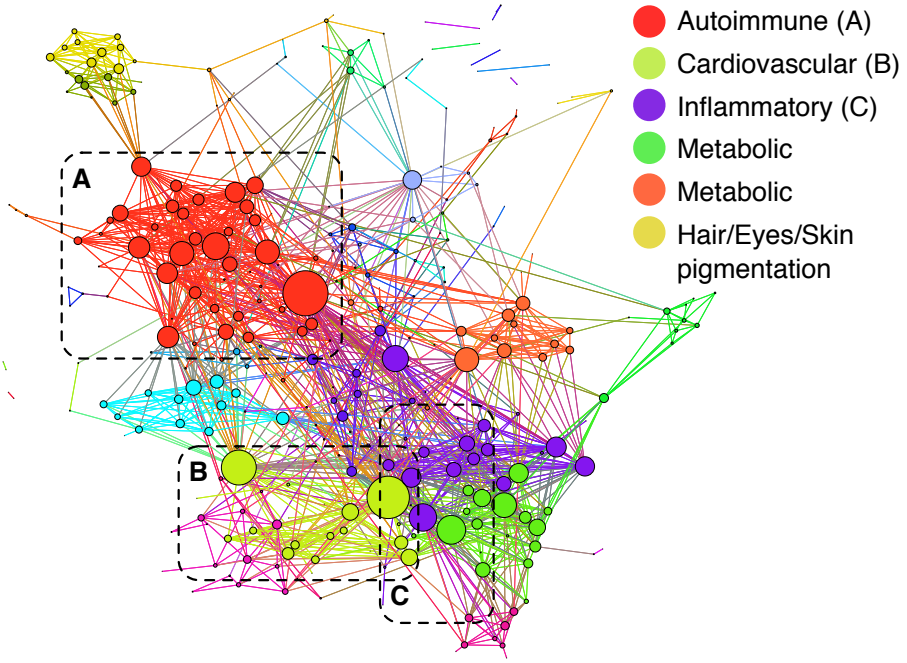

Fig. 4. Network of Phenotypes. Nodes represent phenotypic traits: physical attributes, diseases and behaviors assessed for genetic predisposition through GWAS. Node sizes are proportional to the number of iRAVs associated with the phenotype. The edge weights are based on the number of shared iRAVs normalized by the number of iRAVs associated with the source node. Modules are calculated based on the Louvain method of greedy optimization. Several intuitive modules occur, such as immunological traits or hair and skin pigmentation clusters. The most important modules are shown according to the legend. Chosen modules A, B, and C will be focused on in Fig. 5.

lung cancer and systemic lupus erythematosus (Figs. 5B&C). Consistent with this, it has been shown that lupus patients show an increased risk of lung cancer [4]. Interestingly, the over-the-counter drug cimetidine, found to significantly decrease the lung adenocarcinoma tumor burden compared to untreated controls in mice [19], has also been administered to patients who develop lupus nephritis, an inflammation of the kidney caused by systemic lupus erythematosus, to improve renal function [18]. Finally, the HDN may help uncover biomarkers for diseases. For example, a connection between C-reactive protein (CRP), whose levels in the blood rise in response to inflammation, and Alzheimers disease (AD) is noted (Fig. 5C). Clinically, plasma levels of CRP remain normal in AD patients. However, specific polymorphisms in the regulatory regions of the CRP gene are associated with patients at risk of developing AD [11], different from the SNP shared by the two disease phenotypes. Our results show the implications of disease connectivity, via shared risk-loci, which should help to better understand the shared etiology of linked diseases. This is highlighted by the connection between CRP and AD. CRP was recently found to be a biomarker for AD even though NFκB has

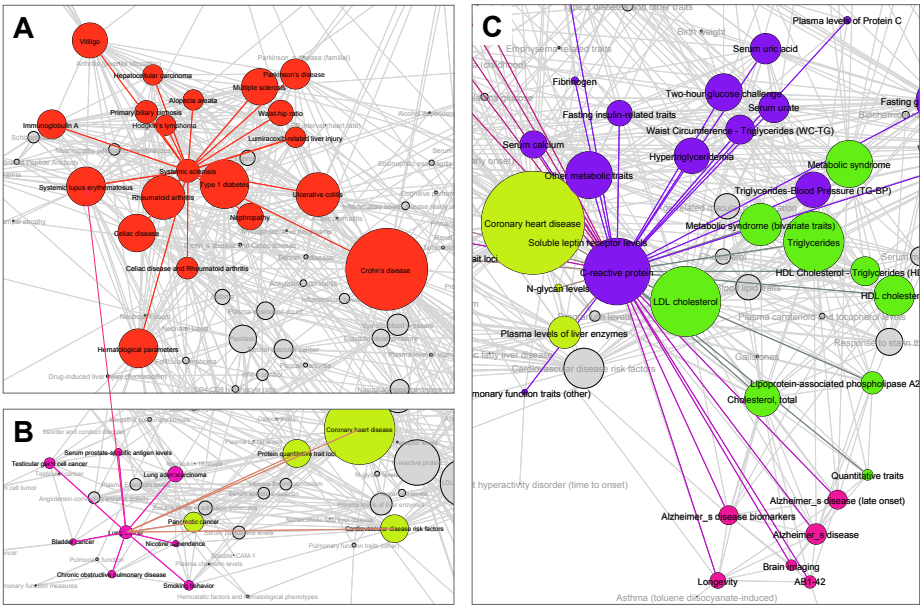

Fig. 5. Network of Phenotypes: In Panel A, traits associated with systemic sclerosis. The connections reflect established associations of systemic sclerosis to immunological diseases and traits clustered within the same module. In Panel B, traits associated with lung cancer. The connection between lung cancer and systemic lupus erythematosus reflects a higher disease risk of lung cancer in cohorts of lupus patients which was only discovered incidentally by previous studies. In Panel C, traits associated with the C reactive protein (CRP). Previous clinical studies have linked CRP to diseases like coronary heart disease and Alzheimers disease.

been long known to play a role in AD development [1]. A connection between the two, however, is seen readily in our network. Additionally, the wealth of known clinical interactions between diseases and traits can be mapped to the risk-associated loci that they share. Hence, this phenotypic network based on genetic predispositions will likely lead to focused biological studies of diseases based on their unique or shared genetic predispositions.

5 Discussion and Future Work

In this study, we successfully build a network of human traits based on about 6,000 genetic variants of over 600 phenotypical traits. We obtain a network that is more comprehensive than that of previous studies by combining GWAS data, which associates PT to raSNPs, with the HapMap project, which links raSNPs to ldSNPs. Our statistical analysis of the network properties places the HPN in the *scale-free* family, showing once more how ubiquitous network structures with heavy-tailed degree distributions really are in the biological, social, and natural networks in general. The automatic classification of phenotypes into “phenotype classes”, using the networks topological modularity

and a standard community detection algorithm, showed very promising results. Indeed, in contrast to what was achieved in previous studies and manual classification, we are able to highlight modules with phenotypes with potentially interesting shared biology, not by arbitrary disease types (i.e. all cancers are classified together regardless of their genetic background). Despite its simplicity, the HPN both confirmed the existence of commonly known phenotype interactions, and also unveiled links that have nevertheless been characterized in recent literature. Because of these findings, we are highly confident that the HPN, and its subsequent revisions, has the potential to become an advantageous clinical tool, both in helping to discover shared biology between PTs, and for possible development of common target drugs. We are currently working on using statistical methods to help us filter out the connections that are genetically and statistically less probable. We would also like to include different datasets, and analyze the overlap of a HDN using genes, pathways, and protein interactions.

Acknowledgments. This work was partially supported by NIH grants R01 EY022300, LM009012, LM010098, AI59694, by the Swiss National Science Foundation grant PBLAP3-136923, by the New Hampshire IDeA Network of Biological Research Excellence (NH-INBRE) with grants from the National Center for Research Resources (5P20RR030360-03) and the National Institute of General Medical Sciences (8P20GM103506-03), and by the Joint Research Program 2011 of the University of Torino and the Compagnia di San Paolo.

References

1. Akiyama, H., Barger, S., Barnum, S., Bradt, B., Bauer, J., Cole, G.M., Cooper, N.R., Eikelenboom, P., Emmerling, M., Fiebich, B.L., Finch, C.E., Frautschy, S., Griffin, W., Hampel, H., Hull, M., Landreth, G., Lue, L.-F., Mrak, R., Mackenzie, I.R., McGeer, P.L., O'Banion, M., Pachter, J., Pasinetti, G., Plata-Salaman, C., Rogers, J., Rydel, R., Shen, Y., Streit, W., Strohmeyer, R., Tooyoma, I., Muiswinkel, F.L.V., Veerhuis, R., Walker, D., Webster, S., Wegrzyniak, B., Wenk, G., Wyss-Coray, T.: Inflammation and alzheimer's disease. Neurobiology of Aging 21(3), 383–421 (2000)
2. Albert, R.: Scale-free networks in cell biology. J. of Cell Science 118, 4947–4957 (2005)
3. Barrenas, F., Chavali, S., Holme, P., Mobini, R., Benson, M.: Network properties of complex human disease genes identified through genome-wide association studies. PLoS ONE 4(11), e8090 (2009)
4. Bin, J., Bernatsky, S., Gordon, C., Boivin, J.-F., Ginzler, E., Gladman, D., Fortin, P.R., Urowitz, M., Manzi, S., Isenberg, D., Rahman, A., Petri, M., Nived, O., Sturfeldt, G., Ramsey-Goldman, R., Clarke, A.E.: Lung cancer in systemic lupus erythematosus. Lung Cancer 56(3), 303–306 (2007)
5. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008(10), P10008+ (2008)
6. Falconer, D.S., Mackay, T.F.C.: Introduction to Quantitative Genetics, 4th edn. Prentice Hall (February 1996)
7. Goh, K.-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabasi, A.-L.: The human disease network. Proceedings of the National Academy of Sciences 104(21), 8685–8690 (2007)

8. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., Manolio, T.A.: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. U S A 106(23), 9362–9367 (2009)
9. The International HapMap Consortium: A haplotype map of the human genome. Nature 437(7063), 1299–1320 (2005)
10. Jobin, C., Sartor, R.B.: Nf-kappab signaling proteins as therapeutic targets for inflammatory bowel diseases. Inflamm. Bowel Dis. 6(3), 206–213 (2000)
11. Kok, E.H., Alanne-Kinnunen, M., Isotalo, K., Luoto, T., Haikonen, S., Goebeler, S., Perola, M., Hurme, M.A., Haapasalo, H., Karhunen, P.J.: Crp gene variation affects early development of alzheimer's disease-related plaques. J. Neuroinflammation 8, 96 (2011)
12. Krum, S.A., Chang, J., Miranda-Carboni, G., Wang, C.-Y.: Novel functions for nfkappab: inhibition of bone formation. Nat. Rev. Rheumatol. 6(10), 607–611 (2010)
13. Li, H., Lee, Y., Chen, J.L., Rebman, E., Li, J., Lussier, Y.A.: Complex-disease networks of trait-associated single-nucleotide polymorphisms (SNPs) unveiled by information theory. Journal of the American Medical Informatics Association: JAMIA 19(2), 295–305 (2012)
14. Loscalzo, J., Barabasi, A.-L.: Systems biology and the future of medicine. Wiley Interdisciplinary Reviews: Systems Biology and Medicine 3(6), 619–627 (2011)
15. McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P.A., Hirschhorn, J.N.: Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat. Rev. Genet. 9(5), 356–369 (2008)
16. Newman, M.: Networks: An Introduction. Oxford University Press, Inc., New York (2010)
17. Osborne, C.K.: Steroid hormone receptors in breast cancer management. Breast Cancer Res. Treat. 51(3), 227–238 (1998)
18. Roubenoff, R., Drew, H., Moyer, M., Petri, M., Whiting-O'Keefe, Q., Hellmann, D.B.: Oral cimetidine improves the accuracy and precision of creatinine clearance in lupus nephritis. Ann. Intern. Med. 113(7), 501–506 (1990)
19. Sirota, M., Dudley, J.T., Kim, J., Chiang, A.P., Morgan, A.A., Sweet-Cordero, A., Sage, J., Butte, A.J.: Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci. Transl. Med. 3(96), 96–477 (2011)
20. Smith, E.P., Boyd, J., Frank, G.R., Takahashi, H., Cohen, R.M., Specker, B., Williams, T.C., Lubahn, D.B., Korach, K.S.: Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man. N. Engl. J. Med. 331(16), 1056–1061 (1994)
21. Suthram, S., Dudley, J.T., Chiang, A.P., Chen, R., Hastie, T.J., Butte, A.J.: Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. PLoS Comput. Biol. 6(2), e1000662 (2010)
22. Watts, D.J.: Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton (1999)
23. Watts, D.J., Strogatz, S.H.: Collective dynamics of "small-world" networks. Nature 393, 440–442 (1998)
24. Wu, X., Jiang, R., Zhang, M.Q., Li, S.: Network-based global inference of human disease genes. Mol. Syst. Biol. 4, 189 (2008)
25. Zhang, X., Cowper-Sal lari, R., Bailey, S.D., Moore, J.H., Lupien, M.: Integrative functional genomics identifies an enhancer looping to the sox9 gene disrupted by the 17q24.3 prostate cancer risk locus. Genome Res. 22(8), 1437–1446 (2012)