Network Analysis in the Investigation of Chronic Respiratory Diseases
From Basics to Application

Diego Diez[1], Alvar Agustí[2], and Craig E. Wheelock[3]

[1]Quantitative Immunology Research Unit, World Premier International Immunology Frontier Research Center, Osaka University, Osaka, Japan; [2]Thorax Institute, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Centro de Investigacion Biomédica en Red de Enfermedades Respiratorias (CIBERES), Madrid, Spain; [3]Department of Medical Biochemistry and Biophysics, Division of Physiological Chemistry II, Karolinska Institutet, Stockholm, Sweden

Abstract

Chronic respiratory diseases are complex multifactorial disorders whose pathogenesis depends on the interplay between host and environmental factors. To fully understand them and to identify novel treatments, a holistic approach that integrates multiple types and levels of clinical and biological data is necessary. Toward this end, the application of systems biology–based strategies, in particular, network analysis, offers great potential. These systems-based approaches rely heavily on computational methods that can be challenging for the nonspecialist. Accordingly, this Pulmonary Perspective: (*1*) outlines the basic concepts of networks in biology and the fundamentals of network analysis, and (*2*) discusses recent applications of network analysis to understand respiratory diseases. The intent of this Perspective is to provide readers with increased understanding of the strengths and weaknesses of network analysis methods as well as their usefulness in addressing research questions involving chronic respiratory diseases.

Keywords: systems biology; systems medicine; protein–protein interaction network; signaling network; gene regulatory network

In many aspects, the lung is a unique organ. It has a complex, multidimensional structure and a rapidly shifting cellular content, and it is directly exposed to a constantly changing environment. There are more than 40 distinct cell types in pulmonary tissue, only 4 of which are considered to be unique to the lung: nonciliated bronchiolar secretory (club [Clara]) cells, type I pneumocytes (squamous cells), type II pneumocytes (great alveolar cells), and alveolar macrophages (1). As a result, chronic respiratory diseases, including asthma, chronic obstructive pulmonary disease (COPD), and pulmonary fibrosis, are complex multifactorial disorders that depend on the interplay between genetic, molecular, cellular, and environmental factors (2–4). It is therefore not surprising that the pathogenesis of many chronic respiratory diseases is still incompletely understood, and curative therapeutics are lacking (4). The traditional reductionist research approaches that focus on a single or small number of molecules (e.g., genes or proteins) cannot address this complexity, and integration of high-dimensional data from multiple clinical and biological levels is the new proposed paradigm to elucidate the pathophysiology of pulmonary diseases (2, 3). An important consequence of this holistic approach is the generation of extremely large datasets (i.e., "omics" datasets, including genomics, transcriptomics, proteomics, and metabolomics) that can provide tens of thousands of data points from a single sample (5). The analysis and interpretation of these enormous datasets has generated distinct challenges, and the accumulation of data *per se* has not increased our insight into the etiology of disease (6).

The fields of systems biology and network medicine have emerged as a response to these difficulties (6–8). Several reviews in respiratory medicine have highlighted the importance of these approaches (9–13), and this strategy has been successfully used in multiple studies (Table 1). Network analysis is a key component of many of these investigations (7), but it can be obfuscating for the nonspecialist. Accordingly, this Pulmonary Perspective: (*1*) outlines the basic concepts of networks and network analysis in biology, and (*2*) discusses its application in chronic respiratory diseases.

(Received in original form March 5, 2014; accepted in final form September 24, 2014)

Supported by Osaka University and JSPS KAKENHI grant 25430182 (D.D.); grants FIS 12/01117, SEPAR 192/2012, and 065/2013, and RESERCAIXA 2012 (A.A.); and the Swedish Heart Lung Foundation and the Center for Allergy Research (C.E.W.).

Correspondence and requests for reprints should be addressed to Diego Diez, Ph.D., Quantitative Immunology Research Unit, WPI Immunology Frontier Research Center (IFReC), 8F Integrated Life Science Building, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan. E-mail: diez@ifrec.osaka-u.ac.jp

Am J Respir Crit Care Med Vol 190, Iss 9, pp 981–988, Nov 1, 2014

Originally Published in Press as DOI: 10.1164/rccm.201403-0421PP on September 25, 2014
Internet address: www.atsjournals.org

Table 1. Applications of Network Analysis to the Study of Respiratory Diseases

Disease	Method Used	Primary Finding	Reference
Asthma	Transcriptome + Bayesian networks	TGF-β_1 is a key regulator of IL13-dependent asthma responses	21
Asthma	Transcriptome + PPI network	PPAR and INF pathways regulate response to glucocorticoids	29
Asthma	SNPs + PPI network	Identified nonsynonymous SNPs in the coding sequences of the toll-like receptor network	39
Asthma	OMIM + transcriptome + PPI network	Suggests that GNB2L1 plays a role as an important signaling mediator in asthma	40
Asthma and smoking	Transcriptome + IPA	TIMP1 and TSBH1 are related to oxidative stress	41
COPD	COPD comorbidities (expert) + PPI	COPD multimorbidities share genes and biological pathways	42
COPD	Transcriptome + mutual information	COPD is associated with failure to regulate bioenergetics pathways	19
COPD	Transcriptome + IPA	There is distinct gender specificity in response to acute smoking and COPD	25
Lung cancer	PPI network	Up-regulated genes tend to be highly connected, contrary to down-regulated genes	28
Multiple (diseasome)	OMIM	Shared genes suggest common origin for diseases	17
Pseudomonas aeruginosa in chronic cystic fibrosis	Transcriptome + metabolome + KEGG	Up-regulation of a pathogen's genes involved in adaptation during disease progression represent potential therapeutic targets	24
Pulmonary hypertension	miRNA transcriptome + target prediction	Identified miRNA-21 as a regulator of disease pathways	22

Definition of abbreviations: COPD = chronic obstructive pulmonary disease; IPA = ingenuity pathway analysis (QIAGEN), commercial software for network/pathway analysis; KEGG = Kyoto Encyclopedia of Genes and Genomes; OMIM = Online Mendelian Inheritance in Man; PPI = protein–protein interactions; SNPs = single-nucleotide polymorphisms; TGF = transforming growth factor.
The diseasome contains information related to several respiratory diseases, including asthma, COPD, and lung cancer.

Cellular Pathways Can Be Represented as Networks

The biochemical and molecular processes occurring in the cell are commonly represented in terms of pathways (Figures 1C–1E). Three general types of pathways can be distinguished: (*1*) metabolic pathways, which represent the reactions involved in the enzymatic transformation of substrates into products; (*2*) signaling pathways, which represent the cascades of interacting proteins and small molecules (ligands and secondary messengers) that enable the transduction of information from outside of the cell; and (*3*) gene regulatory pathways, which represent the regulatory mechanisms that control the transcriptional programs that determine cell identity and phenotypes (including the epigenetic remodeling of chromatin structure and the cooperative binding of transcription factors [TFs] to promoters and enhancers to regulate the expression of target genes).

These pathways were originally described as individual entities, but they are interconnected and undergo cross-talk at multiple levels. Therefore, it is not appropriate to discuss independent pathways, but instead focus should be placed on the system of integrated pathways (i.e., the network of cellular pathways). Accordingly, a cell can be regarded as a system of metabolic, signaling, and gene regulatory pathways involving thousands of molecules, possibly engaged in hundreds of thousands of interactions with other molecules. This complexity enables cellular systems to show extreme robustness and adaptability to different environmental conditions (14) but simultaneously renders them inherently resilient to detailed characterization (i.e., the contribution of each component to the overall system cannot be easily determined). Fortunately, biological pathways are defined in terms of molecules interacting with other molecules, enabling their representation as networks, which can be investigated using network analysis tools (Figures 1C–1E).

In its simplest form, a network is defined as a group of interconnected points, which are composed of a set of nodes and edges that contain information on the system being modeled. Nodes are junction or connection points that represent the variables in the network, and edges portray

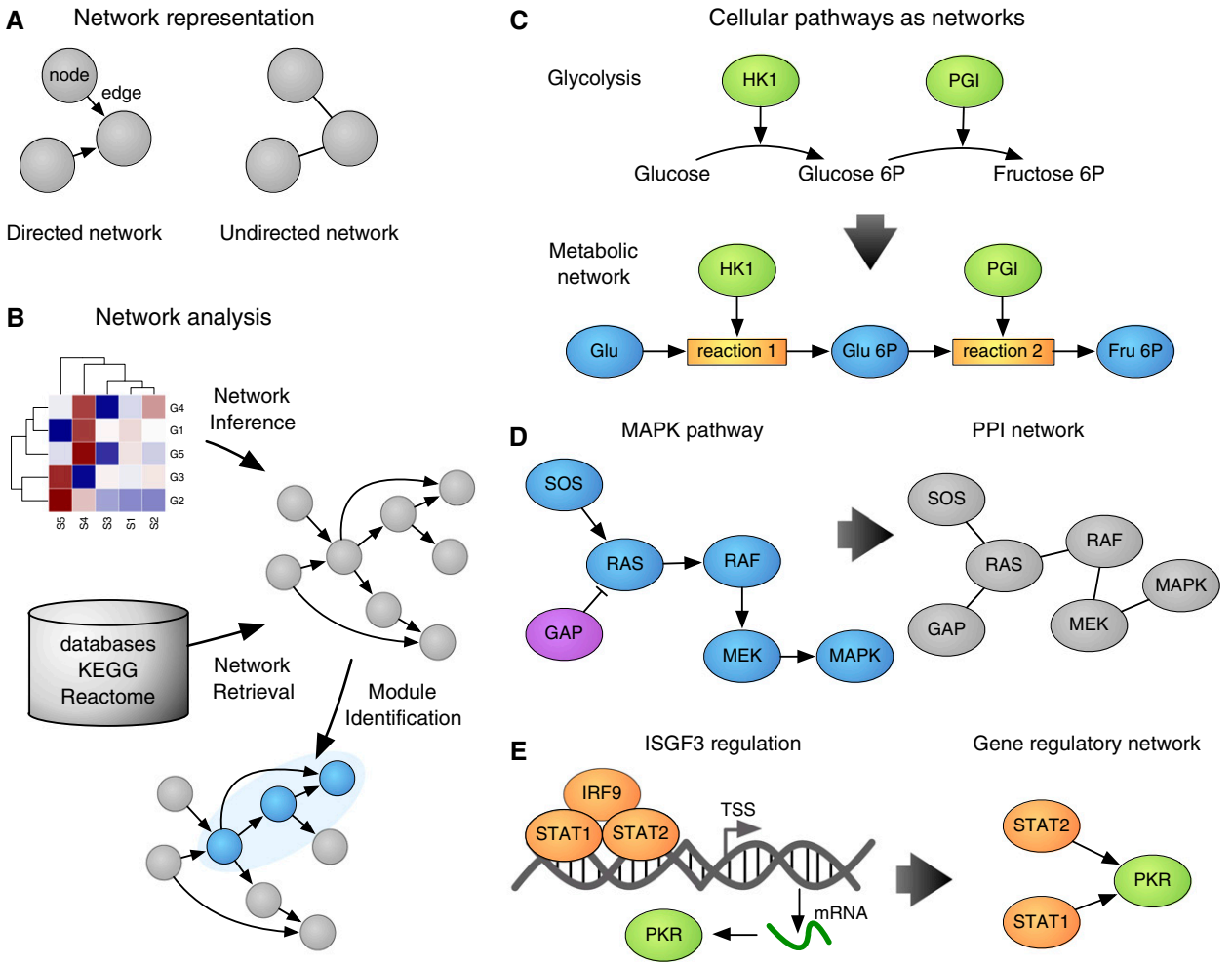

Figure 1. (*A*) Nodes represent entities (e.g., molecules, diseases, conditions) and edges relate nodes to one another. The relations may be directed (*left*) or undirected (*right*). (*B*) Network analysis workflow: the network structure can be learned *de novo* from experimental data or obtained from knowledge databases like Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome. Integrating experimental data enables the identification of disease-related subnetworks (modules). (*C–E*) Cellular pathways can be represented as networks. In metabolic networks (*C*), nodes represent metabolic compounds (Glucose, Glucose-6P, Fructose-6P) and enzymes (HK1 = hexokinase, PGI = phosphoglucoisomerase) connected through a reaction. In signaling networks (*D*), nodes represent proteins (e.g., SOS, RAS, RAF) and edges represent the processes that facilitate (↓) or block (⊥) the flow of information (i.e., signal transduction; e.g., phosphorylation of MEK by RAF1). In gene regulatory networks (*E*), nodes represent transcription factors (TF) (e.g., STAT1 and STAT2) and target genes (e.g., PKR), and edges indicate activation (or repression) of the target gene's expression by the regulating TF. ISGF3 = IFN-stimulated gene factor 3; PPI = protein–protein interactions; TSS = transcription start site.

the relationships among them (Figure 1A). This structure (or layout) can be used to represent multiple types of data from social interactions (e.g., Facebook) and cultural history (15) to cellular signaling (16). This flexibility has fueled its incorporation into research approaches that require the ability to integrate and interrogate large and disparate datasets (e.g., omics-based data). Hence, networks can be used to represent cellular pathways, where the nodes correspond to molecules and the edges to the biochemical actions that molecules perform. Edges can be directed (e.g., protein A phosphorylates protein B, but the opposite is not true) or undirected (e.g., protein A interacts with protein B, and the opposite is also true) (Figure 1A). Nodes and edges in the network convey information about the biological processes occurring in the cell, and this information can be used by computational methods that aim to identify the dynamic behavior of the network (i.e., changes in biochemical and/or physiological activity with time). Accordingly, there is interest in identifying the structure of disease networks and understanding the dynamic effects of perturbations (e.g., gene knock out/knock down, pharmacological intervention, etc.).

Networks can also be used to model other types of information. For example, the so-called human diseasome connects diseases (clinical level) and genes (molecular level) (17). This bipartite network (i.e., a network containing two disjoint sets of nodes: diseases and genes) can produce two different projections: (*1*) the gene network, whose nodes are genes—and two genes are connected if they are associated with the same disease, and (*2*)

the disease network, whose nodes are diseases—and two diseases are connected if they are associated with the same genes (17). Analysis of the disease projection revealed that many diseases share gene connections, suggesting that there are underlying common causes to many disorders (e.g., most cancers are connected through TP53 and PTEN, emphasizing the role that mutations in these genes play in cancer pathogenesis).

Application of Network Analysis to the Investigation of Respiratory Diseases

Network Inference Methods

One of the most common applications of network analysis is network inference, where the structure of the network (that is, the relationships between connecting nodes) is unknown, and the aim is to reconstruct it from experimental data. In other words, the goal is to identify the regulatory relationships and/or interactions between biological components. For example, in signaling networks the objective is to identify which protein regulates which other protein and, in gene regulatory networks, which TFs regulate which target genes.

Experimental determination of cellular pathways is costly and time consuming; hence, methods that automatically reconstruct cellular pathways from experimental omics data have attracted considerable interest. Many approaches have been developed to solve this question and have provided in some cases mechanistic insight. For example, mutual information, a measure of statistical interdependency between two random variables, can be used to reverse-engineer cellular networks (18). Turan and coworkers applied this approach to integrate gene expression, blood cytokine levels, pulmonary gas exchange, and skeletal muscle function in COPD, and identified hypoxia as a potentially critical factor contributing to skeletal muscle dysfunction in these patients (19).

Another approach to network inference involves the use of Bayesian networks, which are a type of probabilistic graphical model where nodes represent conditional probability densities and the network provides the joint probability distribution over the set of variables. Bayesian networks have been applied to successfully reconstruct known signaling and regulatory pathways from gene expression and other types of omics data (16) and also to construct decision systems for diagnosis and treatment of ventilator-associated pneumonia (20). Novershtern and colleagues used Bayesian networks to integrate gene expression profiles in a murine asthma model, which highlighted the role of transforming growth factor-β_1 as a critical regulator of IL13-dependent responses (21).

A general limitation preventing the successful application of network inference approaches to human diseases is the requirement of large datasets generated under different experimental conditions (i.e., system perturbations) to reliably predict interactions among elements of the system. This is a particular obstacle in many clinical studies where sample numbers and the ability to perform perturbations are often limited. A somewhat tangential approach to reconstruct networks is to integrate experimental evidence combined with computational predictions. For example, Parikh and colleagues created a miRNA/target gene network associated with pulmonary hypertension using *in silico* prediction of miRNA targets from miRNA expression data (22). They demonstrated that miRNA-21 represses RhoB and Rho-kinase expression, which is related to decreased angiogenesis and vasodilatation. This approach is also commonly applied to reconstruct gene regulatory networks using TF ChIP-seq data.

Module Identification Methods

The second most common application of network analysis is module identification. A module, or subnetwork, is a distinct subset of functionally related nodes and edges that are expected to physically interact more frequently and show stronger functional dependencies with each other than with other molecules in the network (23). In module identification, the global network is interrogated using a range of topological analysis methods (reviewed in Reference 23) to identify densely interconnected regions associated with some characteristic of interest (e.g., asthma treatment; Figure 1B). These methods are based on the principle that modular structures such as biochemical/signaling cascades or protein complexes exhibit characteristic patterns of interaction. Modules can be further investigated by functional enrichment approaches (e.g., gene ontology enrichment analysis) to evaluate if the molecules implicated in an analysis are functionally relevant as well as discover unexpected shared functions. Module identification can be a powerful approach to identify subphenotypes of disease or unique therapeutic responders and is being incorporated into system medicine–based studies in respiratory disease (5).

Knowledge Databases

The network structure may be inferred from the data, but it may also be obtained from databases that accumulate the current knowledge of cellular pathways. Typically, these databases organize the information distinguishing between metabolic and signaling pathways. Metabolic networks are, by far, the best-characterized biological network, with most of the relevant pathways and relationships (reactions) determined. Several databases provide access to the known metabolic network information in humans and other organisms, including the Kyoto Encyclopedia of Genes and Genomes (KEGG; www.kegg.jp) and the Reactome project (www.reactome.org). For example, Oberhardt and colleagues integrated gene expression data with the metabolic network for the pathogen *Pseudomonas aeruginosa* in the context of chronic cystic fibrosis lung infection and identified pathogen adaptations to disease progression (24). The metabolic network was determined experimentally by flux analysis, and assignment of genes to reactions was accomplished with information from KEGG metabolic pathways and other databases (24).

Signaling networks are relatively uncharacterized, with new pathways being constantly discovered. Current knowledge is cataloged in web databases like KEGG and Reactome. A particularly useful characteristic of the KEGG database is that it contains "disease pathways," which are collections of pathways that contain genes associated with a particular disease (e.g., asthma). This information can be combined with alterations in gene expression to focus on disease-specific alterations. An example is our study that examined the leukocyte transcriptome after exposure to cigarette smoke (25). Results

identified a number of genes (as well as their ontologies and interaction networks) that were differentially expressed exclusively in patients with chronic obstructive pulmonary disease and revealed significant sex differences (25). Unfortunately, the representation of respiratory disease pathways in KEGG is rather limited, only including information on "small cell lung cancer," "non-small cell lung cancer," and "asthma," with the latter being a very naive representation of the complex mechanisms that drive the disease. This highlights a key limitation in the respiratory field, and significant efforts are required to increase the content and functionality of respiratory disease–specific pathways to take full advantage of systems biology approaches.

Protein–Protein Interaction Networks

A key characteristic of signaling networks is that proteins need to interact with other proteins to perform their regulatory functions. Therefore, a way to systematically decipher the cell's signaling networks is to identify the network of protein–protein interactions (PPIs) or interactome (26). Despite some limitations (26), PPI networks have been successfully applied to elucidate biological mechanisms by integrating PPI data with transcriptional changes (27). In these approaches, the statistical significance of the change in expression can be used to score each protein in the PPI network, with higher scores corresponding to genes with lower *P* values. This scored network can be used to identify modules that contain proteins with high scores, which are associated with the cellular responses (Figure 2A) (23). Wachi and colleagues used a combined interactome/transcriptome approach to study the topological features of differentially expressed genes in lung squamous cancer tissue (28). They reported that differentially expressed genes in lung cancer tend to be connected in the PPI network to other differentially expressed genes, suggesting the existence of functional modules of coregulated genes. In a similar fashion, we combined PPI and gene expression data to study the effect of inhaled corticosteroids in patients with asthma treated with fluticasone (29). The identified PPI network module suggested a link between the regulation of gene expression by glucocorticoids, activation of the PPAR pathway (by up-regulation of PPARGC1A and RXRA), and repression of the type I interferon pathway (by down-regulation of STAT1 and IRF9).

Gene Regulatory Networks

Despite the critical role that TFs play in regulating cellular processes and their association with diseases, gene regulatory networks are by far less well characterized (30). Pioneer studies in bacteria and yeast enabled the discovery of essential network motifs (i.e., a recurrent type of subnetwork, like the feedforward loop) that regulate the dynamics of transcriptional regulation (31). However, transcriptional regulation is more complex in higher eukaryotes. Enhancers may be located thousands of nucleotides away from their targets, challenging the identification of functional enhancers/promoters and the pairing of TFs and target genes. Moreover, TFs cooperate with other TFs to form a combinatorial code that regulates the expression of target genes in what are called transcriptional regulatory modules (TRMs). Elucidating these TRMs is key to understanding how TFs regulate transcriptional programs but requires the systematic characterization of many TF binding profiles (32). We recently exploited this idea and developed a new network analysis method (called rTRM) that uses information about PPIs and TF binding to identify TRMs (Figure 2B) (32, 33). The application of rTRM to respiratory disease should help elucidate the role of TFs in driving disease phenotypes.

Network Analysis as an Integrative Platform

Over the past 2 decades there has been a proliferation of technologies capable of measuring a range of molecules in the cell and their associated interactions (5), including probe hybridization technologies

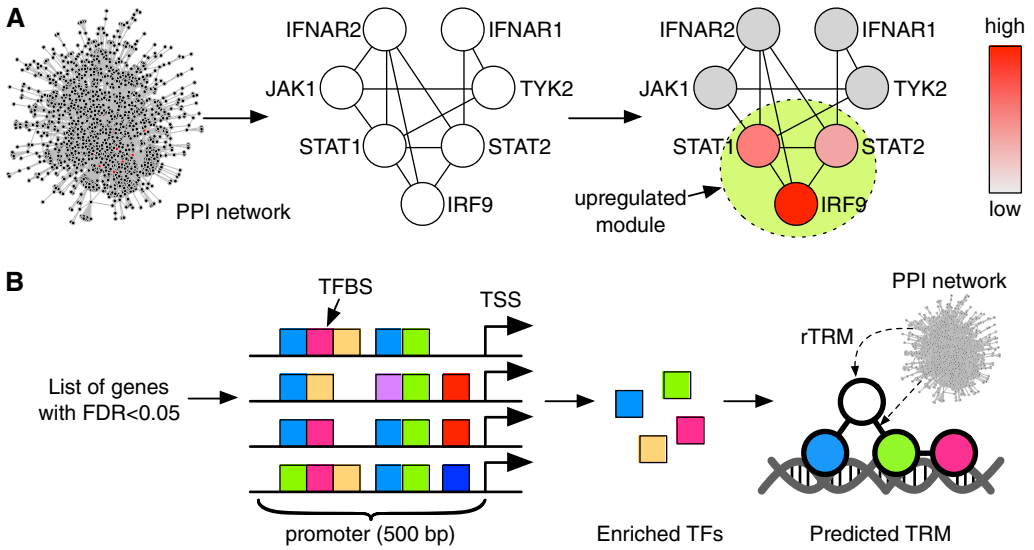

Figure 2. (*A*) Use of protein–protein interaction (PPI) networks to gain insight into signaling networks. *Red nodes* represent proteins with high score (low *P* value). (*B*) Use of PPI networks to investigate transcriptional regulatory networks. FDR = false discovery rate; TF = transcription factor; TFBS = TF binding sites; TSS = transcription start site; TRM = transcriptional regulatory modules.

Table 2. Selected Software Tools for Network Analysis

Name	Platform	Purpose	Ref
bayelviraApp	Cytoscape	Bayesian network inference	43
jActiveModules	Cytoscape	Identification of network modules	27
CyTargetLinker	Cytoscape	Link regulatory (transcription factor) information	44
DisGeNET	Cytoscape	Query disease–gene networks	45
minet	Bioconductor	Inference of networks using mutual information	46
birta	Bioconductor	Bayesian network inference	47
BioNet	Bioconductor	Identification of network modules	48
rTRM	Bioconductor	Identification of transcriptional regulatory modules	32
Cyrface	Cytoscape	Interface between Cytoscape and R	49
RCytoscape	Bioconductor	Interface between Cytoscape and R	50

Selected tools from the Cytoscape (http://apps.cytoscape.org) and Bioconductor (http://www.bioconductor.org/packages/release/) packages for the analysis of biological networks. Additional information for other software tools for network analysis in R can be found at CRAN (http://www.r-project.org).

(e.g., microarrays, ChIP-on-chip), sequencing technologies (e.g., RNA-seq, ChIP-seq), and mass spectrometry (e.g., proteome, phosphoproteome, metabolome, glycome, lipidome). These approaches have resulted in a substantial increase in the available data to investigate respiratory diseases (5). Because developments in these technologies will certainly continue to increase, the amount, type, and structure of data available for systems medicine–based studies will grow concomitantly. Accordingly, a major challenge is how to integrate these datasets into coherent disease models. Network analysis provides a framework to incorporate this information. For example, nodes can integrate information on different measurements made by different technologies on the same molecule. Likewise, nodes representing diseases or phenotypes can also be integrated into the network via connecting them to other nodes (e.g., diseases connected to genes in the diseasome network [17, 34]). However, a difficulty of this approach is that increasing complexity in the network model results in commensurate increase in computational complexity (i.e., increased number of parameters). Furthermore, there is generally limited availability of experimental data measuring the effect(s) of perturbations (e.g., disease state, knock out, knock down, etc.), over time, which is necessary to reliably estimate the dynamic behavior of the network. However, as the availability of experimental data on biochemical and pathophysiological interactions in the lung continues to accumulate, our ability to apply network methodologies will expand.

Software Tools for Network Analysis

There is a plethora of software tools for network analysis, which cannot be succinctly reviewed herein, and the interested reader is referred to recent reviews (35, 36). We highlight here two popular projects, Cytoscape and Bioconductor, which provide a variety of free tools for network analysis. Both resources are well-established projects, likely ensuring their long-term continuity and adaption to new technologies. It is expected that these projects will be the primary tools used in the applications of network analysis to explore pulmonary diseases.

Cytoscape (http://cytoscape.org) provides a point-and-click user-friendly interface (37). The basic functionality includes reading, annotating, and manipulating networks. The software environment can accept multiple types of data in a range of different formats, which can then be incorporated into network generation. It is also possible to import network information from biological databases, including KEGG, Reactome, and BioGRID. Additional capabilities are available as plug-ins contributed by the research community (Table 2).

The Bioconductor project (http://bioconductor.org [38]) implements methods for the analysis of biological data for the R platform (http://r-project.org). Although working in R may be challenging for the uninitiated user, its flexibility and powerful statistical capabilities, combined with the large collection of contributed packages, makes R one of the most popular statistical software packages. There are many R packages focusing on biological applications and networks. Most of these packages can be found at the CRAN repository (http://cran.r-project.org) and at Bioconductor (Table 2). Finally, a number of packages allow communication between Cytoscape and R, providing the best of each platform (Table 2).

Conclusions

The increasing accumulation of omics-based datasets describing cellular mechanisms and their relation to clinical phenotypes holds the potential to elucidate mechanisms of complex lung diseases. However, this information cannot be easily analyzed or understood using classical research approaches, and novel systems-based integrative methods are needed. Network analysis provides a starting point for these new analytical approaches, which needs to be combined with rigorous clinical phenotyping to provide quantitative patient profiling. Yet, although promising, there are a number of obstacles in the widespread adoption of these research approaches, including: (*1*) insufficient methods for data integration and network generation/interrogation, especially for the nonspecialist; (*2*) paucity of databases with sufficient biological content, particularly in relation to pulmonary diseases; and (*3*) scarcity of interactions between clinical and informatics researchers. We believe that these limitations will be addressed in the near future, and network analyses will become a common approach in systems medicine–based studies of human respiratory diseases. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1. Sorokin SP. The cells of the lungs. In: Nettesheim P, Hanna MG Jr, Deatherage JW Jr, editors. Morphology of experimental respiratory carcinogenesis: proceedings of a Biology Division, Oak Ridge National Laboratory Conference. Gatlinburg, TN: National Technical Information Service; 1970. pp. 3–43.
2. Agusti A, Sobradillo P, Celli B. Addressing the complexity of chronic obstructive pulmonary disease: from phenotypes and biomarkers to scale-free networks, systems biology, and P4 medicine. *Am J Respir Crit Care Med* 2011;183:1129–1137.
3. Faner R, Rojas M, Macnee W, Agustí A. Abnormal lung aging in chronic obstructive pulmonary disease and idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186:306–313.
4. Wu W, Kaminski N. Chronic lung diseases. *Wiley Interdiscip Rev Syst Biol Med* 2009;1:298–308.
5. Wheelock CE, Goss VM, Balgoma D, Nicholas B, Brandsma J, Skipp PJ, Snowden S, Burg D, D'Amico A, Horvath I, *et al*. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. *Eur Respir J* 2013;42:802–825.
6. Hood L, Balling R, Auffray C. Revolutionizing medicine in the 21st century through systems approaches. *Biotechnol J* 2012;7: 992–1001.
7. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;12:56–68.
8. Bousquet J, Anto JM, Sterk PJ, Adcock IM, Chung KF, Roca J, Agusti A, Brightling C, Cambon-Thomsen A, Cesario A, *et al*. Systems medicine and integrated care to combat chronic noncommunicable diseases. *Genome Med* 2011;3:43.
9. Studer SM, Kaminski N. Towards systems biology of human pulmonary fibrosis. *Proc Am Thorac Soc* 2007;4:85–91.
10. Ahmad F, Champion HC, Kaminski N. Toward systems biology of pulmonary hypertension. *Circulation* 2012;125:1477–1479.
11. Maher TM. Beyond the diagnosis of idiopathic pulmonary fibrosis; the growing role of systems biology and stratified medicine. *Curr Opin Pulm Med* 2013;19:460–465.
12. Dahlin A, Tantisira KG. Integrative systems biology approaches in asthma pharmacogenomics. *Pharmacogenomics* 2012;13: 1387–1404.
13. Faner R, Cruz T, López-Giraldo A, Agustí A. Network medicine, multimorbidity and the lung in the elderly. *Eur Respir J* 2014;44: 775–788.
14. Kitano H. Biological robustness. *Nat Rev Genet* 2004;5:826–837.
15. Schich M, Song C, Ahn YY, Mirsky A, Martino M, Barabási AL, Helbing D. Quantitative social science. A network framework of cultural history. *Science* 2014;345:558–562.
16. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 2005;308:523–529.
17. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. *Proc Natl Acad Sci USA* 2007;104: 8685–8690.
18. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. *Nat Genet* 2005;37:382–390.
19. Turan N, Kalko S, Stincone A, Clarke K, Sabah A, Howlett K, Curnow SJ, Rodriguez DA, Cascante M, O'Neill L, *et al*. A systems biology approach identifies molecular networks defining skeletal muscle abnormalities in chronic obstructive pulmonary disease. *PLOS Comput Biol* 2011;7:e1002129.
20. Visscher S, Lucas PJ, Schurink CA, Bonten MJ. Modelling treatment effects in a clinical Bayesian network using Boolean threshold functions. *Artif Intell Med* 2009;46:251–266.
21. Novershtern N, Itzhaki Z, Manor O, Friedman N, Kaminski N. A functional and regulatory map of asthma. *Am J Respir Cell Mol Biol* 2008;38:324–336.
22. Parikh VN, Jin RC, Rabello S, Gulbahce N, White K, Hale A, Cottrill KA, Shaik RS, Waxman AB, Zhang YY, *et al*. MicroRNA-21 integrates pathogenic signaling to control pulmonary hypertension: results of a network bioinformatics approach. *Circulation* 2012;125: 1520–1532.
23. Mitra K, Carvunis AR, Ramesh SK, Ideker T. Integrative approaches for finding modular structure in biological networks. *Nat Rev Genet* 2013;14:719–732.
24. Oberhardt MA, Goldberg JB, Hogardt M, Papin JA. Metabolic network analysis of Pseudomonas aeruginosa during chronic cystic fibrosis lung infection. *J Bacteriol* 2010;192:5534–5548.
25. Faner R, Gonzalez N, Cruz T, Kalko SG, Agustí A. Systemic inflammatory response to smoking in chronic obstructive pulmonary disease: evidence of a gender effect. *PLoS ONE* 2014;9:e97491.
26. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell* 2011;144:986–998.
27. Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 2002;18:S233–S240.
28. Wachi S, Yoneda K, Wu R. Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* 2005;21:4205–4208.
29. Diez D, Goto S, Fahy JV, Erle DJ, Woodruff PG, Wheelock AM, Wheelock CE. Network analysis identifies a putative role for the PPAR and type 1 interferon pathways in glucocorticoid actions in asthmatics. *BMC Med Genomics* 2012;5:27.
30. Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* 2012;13:613–626.
31. Alon U. Network motifs: theory and experimental approaches. *Nat Rev Genet* 2007;8:450–461.
32. Diez D, Hutchins AP, Miranda-Saavedra D. Systematic identification of transcriptional regulatory modules from protein-protein interaction networks. *Nucleic Acids Res* 2014;42:e6.
33. Hutchins AP, Diez D, Takahashi Y, Ahmad S, Jauch R, Tremblay ML, Miranda-Saavedra D. Distinct transcriptional regulatory modules underlie STAT3's cell type-independent and cell type-specific functions. *Nucleic Acids Res* 2013;41:2155–2170.
34. Menche J, Sharma A, Cho MH, Mayer RJ, Rennard SI, Celli B, Miller BE, Locantore N, Tal-Singer R, Ghosh S, *et al*. A diVIsive Shuffling Approach (VIStA) for gene expression analysis to identify subtypes in chronic obstructive pulmonary disease. *BMC Syst Biol* 2014;8:S8.
35. Thomas S, Bonchev D. A survey of current software for network analysis in molecular biology. *Hum Genomics* 2010;4:353–360.
36. Diez D, Wheelock AM, Goto S, Haeggström JZ, Paulsson-Berne G, Hansson GK, Hedin U, Gabrielsen A, Wheelock CE. The use of network analyses for elucidating mechanisms in cardiovascular disease. *Mol Biosyst* 2010;6:289–304.
37. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–2504.
38. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al*. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80.
39. Renkonen J, Joenvaara S, Parviainen V, Mattila P, Renkonen R. Network analysis of single nucleotide polymorphisms in asthma. *J Asthma Allergy* 2010;3:177–186.
40. Hwang S, Son SW, Kim SC, Kim YJ, Jeong H, Lee D. A protein interaction network associated with asthma. *J Theor Biol* 2008;252:722–731.
41. Freishtat RJ, Benton AS, Watson AM, Wang Z, Rose MC, Hoffman EP. Delineation of a gene network underlying the pulmonary response to oxidative stress in asthma. *J Investig Med* 2009;57:756–764.
42. Grosdidier S, Ferrer A, Faner R, Pinero J, Roca J, Cosio B, Agusti A, Gea J, Sanz F, Furlong LI. Network medicine analysis of COPD multimorbidities. *Respir Res* 2014;15:111.
43. Elvira Consortium. Elvira: an environment for creating and using probabilistic graphical models. In: Gámez JA, Salmerón A, editors. Proceedings of the First European Workshop on Probabilistic Graphical Models (PGM'02). Cuenca, Spain: Electronic Proceedings; 2002. pp. 222–230.
44. Kutmon M, Kelder T, Mandaviya P, Evelo CT, Coort SL. CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis. *PLoS ONE* 2013;8:e82160.

45. Bauer-Mehren A, Rautschka M, Sanz F, Furlong LI. DisGeNET: a Cytoscape plugin to visualize, integrate, search and analyze gene-disease networks. *Bioinformatics* 2010;26:2924–2926.
46. Meyer PE, Lafitte F, Bontempi G. minet: a R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics* 2008;9:461.
47. Zacher B, Abnaof K, Gade S, Younesi E, Tresch A, Fröhlich H. Joint Bayesian inference of condition-specific miRNA and transcription factor activities from combined gene and microRNA expression data. *Bioinformatics* 2012;28:1714–1720.
48. Beisser D, Klau GW, Dandekar T, Müller T, Dittrich MT. BioNet: an R-Package for the functional analysis of biological networks. *Bioinformatics* 2010;26:1129–1130.
49. Gonçalves E, Saez-Rodriguez J. Cyrface: an interface from Cytoscape to R that provides a user interface to R packages. *F1000 Res* 2013;2:192.
50. Shannon PT, Grimes M, Kutlu B, Bot JJ, Galas DJ. RCytoscape: tools for exploratory network analysis. *BMC Bioinformatics* 2013;14:217.