HIGHLIGHTED TOPIC | Emergent Behavior in Lung Structure and Function

Complex systems in pulmonary medicine: a systems biology approach to lung disease

David A. Kaminsky,1 Charles G. Irvin,1 and Peter J. Sterk2

1Vermont Lung Center, Department of Medicine, Division of Pulmonary and Critical Care Medicine, University of Vermont College of Medicine, Burlington, Vermont; and 2Department of Respiratory Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Submitted 5 November 2010; accepted in final form 22 December 2010

Kaminsky DA, Irvin CG, Sterk PJ. Complex systems in pulmonary medicine: a systems biology approach to lung disease. J Appl Physiol 110: 1716–1722, 2011. First published December 23, 2010; doi:10.1152/japplphysiol.01310.2010.—The lung is a highly complex organ that can only be understood by integrating the many aspects of its structure and function into a comprehensive view. Such a view is provided by a systems biology approach, whereby the many layers of complexity, from the molecular genetic, to the cellular, to the tissue, to the whole organ, and finally to the whole body, are synthesized into a working model of understanding. The systems biology approach therefore relies on the expertise of many disciplines, including genomics, proteomics, metabolomics, physiomics, and, ultimately, clinical medicine. The overall structure and functioning of the lung cannot be predicted from studying any one of these systems in isolation, and so this approach highlights the importance of emergence as the fundamental feature of systems biology. In this paper, we will provide an overview of a systems biology approach to lung disease by briefly reviewing the advances made at many of these levels, with special emphasis on recent work done in the realm of pulmonary physiology and the analysis of clinical phenotypes.

emergence; complexity; biological network; physiome; clinical phenotype

THE LUNG is a remarkably complex organ whose range of function varies from the molecular (e.g., gas exchange) to the whole organ level (e.g., bulk flow of air). An understanding of lung health and disease must therefore necessarily incorporate information from across this broad scale of organ structure and function (Fig. 1). This information has to come from integrative physiology and is currently best obtained by a quantitative systems biology approach (6, 70).

In systems biology, the approach to understanding a biological system involves integrating information from all levels of structure and function of the system (1, 36, 37). This is, in fact, what respiratory scientists do when they describe lung function in terms of the forced expiratory volume in 1 s (FEV1), which is a global, robust measure of function, such as in chronic obstructive pulmonary disease (COPD) (67). This process differs from the traditional “reductionist” approach, which involves taking a system apart in order to identify and understand its component parts. For example, consider the complexities involved in the study of the transcription factor nuclear factor kappa B (NFκB) and its role in inflammation (68). While this reductionist process is critical to understanding basic scientific and biological principles, complex biological systems reveal a functional behavior that cannot be predicted just by the sum of their parts. Such a property is known as “emergence” and is the key principle in systems biology (1, 37, 40). Other important principles that are cornerstones of this approach to the complex, network structure of biological systems include “robustness” and “modularity” (1, 37). Robustness confers stability to a system in response to perturbation and is generally characterized by the presence of feedback loops, redundancy, and structural stability (3). Modularity relates to how multiple components of a system function in such a manner as to lead to similar outcomes, providing robustness for the system and allowing it to evolve (3). Together, these properties of emergence, robustness, and modularity comprise the system structure and function, and intrinsic methods for control and design, that are needed to allow a complex biological system to adapt, function, and evolve in response to variations and perturbations in its environment (37).

Given the lung’s highly complex structure and function, a systems biology approach to understanding the lung is an ideal application of the method. Applying systems biology to the study of lung disease involves the use of many resources (59). These include a multidisciplinary team of scientists, from biologists to mathematicians and clinicians, who can apply their unique knowledge and skills and communicate with one another across disciplines in order to acquire data sets (fre-
Recently large in size) and discover links between the various parts of a complex system; the application of an iterative approach to discovery and problem solving, by using mathematical models based on experimental data, and then refining the models as they are tested and new data are obtained; and utilization of advanced mathematical and computational methods to handle large quantities of data, statistically analyze the data, and apply an iterative modeling approach (Fig. 2).

Most papers dealing with systems biology focus on the utilization of high-throughput genomic data to understand patterns of gene expression in relation to phenotypic features of various diseases. Other “-omic” approaches also include proteomics and metabolomics, related to protein expression and metabolic biomarkers, respectively. While these approaches are obviously important components of systems biology, the physiological (physiomic) and clinical aspects of the disease are often not emphasized. Indeed, the IUPS human physiome project was specifically designed to model the human body through advanced computational modeling that integrates data from the cell, tissue, organ, and organ system (33). This more comprehensive approach has been dubbed the “renaissance of physiology,” as traditional physiological approaches must now incorporate data and principles from the many fields of “-omics” that are used in the systems biology approach (58). Applying this integrative strategy to medicine in general defines what has become known as “systems medicine” (7). The purpose of the present paper is to describe a systems biology approach to lung disease that incorporates not only genomic information but also physiological and clinical data that contribute to the phenotypic expressions of disease.

**SYSTEMS BIOLOGY APPROACH USING HIGH-THROUGHPUT DATA**

The key steps involved in a systems biology approach include data acquisition, model construction, disease simulation, model perturbation, theory formation and prediction, and experimental validation of theory (37). According to Wu and Kaminski, however, “... the majority of the studies in chronic lung diseases is still in the first step of systems biology research” (70); i.e., data acquisition. Most data acquisition occurs through the utilization of high-throughput methods involving microarrays of gene expression or more recently RNA sequencing techniques. These data reveal information...
about the complex composition of mRNAs expressed in the tissue or cell of interest. The systems approach to analyzing these data is to look for transcriptional profiles of different phenotypes of tissue in order to associate these profiles with the phenotype. Therefore, this strategy is essentially unbiased, meant to discover previously unknown associations and thereby generate novel hypotheses. This approach has been taken for a number of lung diseases, including idiopathic pulmonary fibrosis (IPF), asthma, COPD, and pulmonary hypertension (70). As an example, a systems approach has lead to important new insights in IPF (70). Traditionally thought of as an inflammatory disease, clinicians know far too well that anti-inflammatory therapy has little effect. Microarrays have shown, however, that genes involved in wound healing and excessive deposition of extracellular matrix are upregulated in this disease. In particular, these gene profiles are distinct from the proinflammatory genes that are active in a disease like hypersensitivity pneumonitis, which clearly involves inflammation and responds to anti-inflammatory therapy (70). Similarly, microarrays of tissue from patients with asthma reveal upregulation of genes involved in the immune response, such as antigen processing, cytokine production, and T-cell regulation (70). In COPD, genes involved in stress-related activation have been identified (70), and recently, genes associated with degradation around small airways have been found (27). Functional genomic profiling has been used extensively in the area of non-small cell lung cancer. A recent review (56) highlights many of the findings in this area, among which has been the recognition of alterations in epidermal growth factor receptor (EGFR) and subsequent development of targeted therapy with tyrosine kinase inhibitors. Recently, Goto and colleagues (28) discovered a unique epigenetic profile of genes in malignant pleural mesothelioma, which was distinct enough that it may serve as a diagnostic marker of disease. Other examples include the upregulation of genes involved in inflammation and stress response in pulmonary hypertension (13), and genes associated with various aspects of lung structure and inflammation associated with susceptibility to acute lung injury (20). Thus high-throughput analyses of gene expression have yielded unique insights into the patterns of biological function associated with distinct lung diseases.

Regulation of posttranscriptional gene expression is also governed by microRNAs (miRNA). Since these molecules will alter ultimate gene expression, molecular methods to characterize patterns of miRNAs are also involved in a systems approach to lung disease. For example, Tan and colleagues (60) have shown that miRNAs are specifically targeted to HLA-G in patients at risk of asthma. Other aspects of epigenetics may also be involved in determining risk and severity of asthma and therefore be useful in classifying phenotype and designing targeted therapies (32).

The data from microarrays can also be used in genomewide association studies (31), where a common approach is to search for single-nucleotide polymorphisms (SNPs) within a population. This approach is particularly used to elucidate pharmacogenetic factors associated with treatment responses, such as, in asthma, the bronchodilator response (17, 69) or the response to anti-leukotriene therapy (46). However, focusing on all the polymorphisms associated with a functional phenotype, rather than just isolated ones, is a more powerful method that a systems biology approach can bring to genomewide association studies (66). A recent example of this approach was described by Moffatt and colleagues (44), who demonstrated that a number of polymorphisms are associated with asthma and, as a whole, implicate an interactive role between epithelial damage, adaptive immunity, and airway inflammation in the pathogenesis of the disease.

Protein expression profiles, or proteomics, are also used in a systems biology approach to lung disease. Recent examples include IPF, asthma, and pulmonary hypertension (70). Since proteins are measurable products in blood and other body fluids, this approach is key to searching for relevant biomarkers...
of disease. Notably, serum protein expression profiles have recently been shown to be highly successful in discriminating the chronic airway diseases asthma, COPD and cystic fibrosis (26). In addition, the analysis of global metabolic profiles, so-called metabolomics, of substances in exhaled breath condensate by NMR spectroscopy, reveals airway biochemical “fingerprints” (14) of oxidized and acetylated compounds in patients with asthma. Furthermore, “breathprints” of volatile organic compounds by electronic noses and/or mass spectrometry can distinguish between patients with asthma and those with COPD (18), and between patients with lung cancer and patients with other malignancies (51). Even though these findings require external validation according to the STARD Guidelines (Standards for the Reporting of Diagnostic Accuracy Studies, www.stard-statement.org), they do suggest that proteomic and metabolomic assessments can be extremely helpful in discriminating patients with lung diseases.

While all of these “-omic” approaches to understanding lung disease are necessary, they are not sufficient. Only through integration of this information through unbiased statistical methods can we derive a complex systems view of lung disease and produce insights that will enhance our understanding of disease classification and pathophysiology (39). For example, from mouse models of asthma, a “functional and regulatory map” constructed using a synthesis of microarray data sets has revealed distinct patterns of response to treatment and led to new insight into the role of IL-13 and transforming growth factor-beta (TGF-B) as key regulators of asthma (48). However, at this time, further integration of molecular “-omics” data into a comprehensive, systems view of human lung disease is still in its infancy (70).

SYSTEMS BIOLOGY APPROACH THROUGH LUNG MECHANICAL BEHAVIOR: TIME SERIES

An essential characteristic of biological systems is their temporal behavior. In-built feedback loops are creating a dynamic homeostasis in physiology, which is better labeled as “homeokinesis” (40). Thus far the dynamics of the respiratory system have not been studied using molecular profiles, but there is increasing evidence of highly characteristic dynamic behavior in relation to disease from measures of respiratory microscopic structure and in vitro and in vivo airway function (42).

As the lung is a highly mechanical structure, it makes sense that searching for emergent mechanical properties would be an important systems biology strategy in lung disease. This is especially true in a disease like asthma that is characterized by temporal variability, which might explain why the reductionist approach to understanding this disease has been limited. A recent example of the failure of the reductionist approach in lung disease is found in the work by LaPrad and colleagues (38), who showed that, contrary to expectations, oscillations of an airway smooth muscle while intact within the airway wall did not attenuate its ability to constrict in response to acetylcholine, unlike the behavior of isolated smooth muscle strips. As pointed out by Bates, “...we still cannot tell a priori how behavior will cross length scales in biological systems” (8).

Bates and colleagues (9) have recently published an elegant example of how emergent properties govern the behavior of biological networks. Using an elastic network model of the lung parenchyma, these authors were able to demonstrate the process of percolation, an emergent property known to be involved in many natural processes, such as forest fires and infectious disease epidemics. Percolation describes the process by which physical events are transmitted across a network of elements. In the lung, these physical events might include the laying down of collagen in the case of IPF, or the destruction of the lung parenchyma in the case of emphysema. Initially, these events might progress in a gradual manner without much overall change in the bulk mechanical properties of the lung. However, at some point, known as the percolation threshold, these events link together across the entire lung to abruptly change its bulk mechanical properties. Clinically, this process may be manifested as gradual progression of disease, followed by acceleration of the disease process when a percolation threshold is reached. Indeed, in IPF an initial gradual progression of disease may be followed by an acute exacerbation of IPF (35). Similarly, in emphysema, radiographic evidence of disease may exist before significant changes in lung function occur (49), but progressive emphysema and accelerated loss of lung function typically follow after a period of time (19, 43). In both cases, radiographic and pathologic data support the concept of percolation as it is predicted to occur in these two diseases (8). Although greatly simplified, this modeling concept may have important clinical implications for the way in which progression of lung disease occurs and how it can be monitored by measuring various aspects of structure by imaging or biopsy.

Another example of applying a systems biology approach to lung disease is in predicting asthma exacerbations based on underlying fluctuation of airway function. Asthma is a complex, chronic disease (5, 64). The complexity of asthma arises from the nonlinear interaction of inflammatory, immunological, physiological, mechanical, and environmental factors, all within a context of genetic background and unique, personality traits and behavior (30, 63). It is no wonder that no one molecule, and no one therapy, can “cure” asthma or apply equally across individuals. Likewise, it is difficult to isolate clinical factors that may predict poor asthma control (41). However, asthma can be approached as any complex system by using a systems biology approach to understand the interaction of these many factors (6, 21).

One of the manifestations of these complex interactions is the fractal nature of lung function, which no doubt arises, in part, from the fractal structure of the airway tree (10). Such a structure is characterized by anatomic similarities at different length and time scales, a so-called scale-free network (4). A unique feature of such a network is that, while it remains robust to small perturbations, it can suddenly compensate with an “avalanche” of heterogeneous airway narrowing that leads to reductions in airway function and a loss of asthma control (4). However, because the fractal nature of lung function is scale free, it can be statistically analyzed at different time scales, such that previous events contain information that leads to prediction of future events (22, 23).
the correlation, and whether future risk of exacerbation could be predicted from fluctuations in PEF. The authors found that, based on computer modeling, the regular use of short-acting beta-agonists, but not long-acting beta-agonists, would result in a more random, and hence less predictable, fluctuation in lung function and an increased risk of future exacerbations (22, 23). In addition, they found that their model allowed baseline fluctuation in PEF to predict the probability of future exacerbations. Such computer modeling was applied in a subsequent study (61) that demonstrated the utility of baseline fluctuation in PEF to predict future asthma symptoms and treatment response to beta-agonists. These studies certainly shed new light on the importance of monitoring of asthma by PEF, which, despite recommendations by international guidelines (25), is notoriously underutilized (62). Interestingly, a recent study involving user-friendly, electronic monitoring in the context of strong clinical support was able to demonstrate markedly improved adherence to PEF monitoring by patients with asthma (53). Such electronic devices could potentially be adapted to collect and analyze time series data.

Another view on temporal behavior of the lung is provided by the short-term fluctuations of respiratory impedance over minutes, rather than days, as measured by the forced oscillation technique. Even though the first fluctuation analyses of respiratory impedance were not consistent in discriminating asthma from controls (16, 52), a recent analysis using a nonlinear, distance-based deterministic model (attractor analysis) did show adequate distinction between patients with asthma and COPD (47). This suggests that the dynamics of the respiratory system provides relevant information in relation to airways disease.

**SYSTEMS BIOLOGY APPROACH THROUGH CLINICAL PHENOTYPES**

Another scale at which to apply a systems biology approach to lung disease is at the clinical level (6, 7). Examining differing patterns of clinical expression of disease can be done by complex statistical approaches such as factor analysis (55) or cluster analysis (45). For example, cluster analysis is a mathematical method that quantifies the similarity between individuals in a population based on specified variables, and groups individuals together such that individuals within the same cluster are highly similar, and individuals in different clusters are very different. Two recent examples applied to asthma illustrate this principle. In the study by Haldar and colleagues (30, 371 patients with asthma were found to group into five phenotypic types that corresponded to specific responses to therapy. These included patients characterized by varying degrees of eosinophilic inflammation, symptoms, and body mass index. Similarly, Moore and colleagues (45) found five clusters of asthma phenotypes that were distinguished by lung function, atopy, age of onset, and use of corticosteroids.

Cluster analysis has also been applied to 175 random subjects with respiratory symptoms or airflow obstruction by spirometry and, interestingly, also found five distinct phenotypes, in this case related to variability of airflow obstruction, atopy, eosinophilic inflammation, emphysema, and chronic bronchitis (63). A similar method, latent class analysis, was able to distinguish five phenotypes of children with wheeze and cough (57). In COPD, Paoletti and colleagues (50) recently applied multiple advanced methods to patients and found phenotype clusters corresponding to the classic types involving bronchitis vs. emphysema. Hurst and colleagues (34) demonstrated that exacerbations of COPD cluster in time, revealing a high-risk period for recurrent exacerbation in the 8-wk period following the initial exacerbation.

Why should we reorganize our thinking about the phenotypic expression of lung disease? The answer lies in the recognition that traditional classifications of phenotype do not appear to relate to mechanisms of disease or to prognostic information or therapeutic outcomes. For example, using FEV$_1$ to classify severity of disease in either COPD or asthma is inadequate, as shown by the wide distribution of symptoms and comorbidities across GOLD stages in the ECLIPSE cohort of patients with COPD (2), or the variable distribution of clinical attributes across degrees of lung function abnormality in a cohort of patients with severe asthma (45). Indeed, in many cases, the use of FEV$_1$ and other traditional clinical and physiological features fails to even distinguish asthma from COPD (24). This inability to better phenotype patients may be one of the reasons that clinical trials fail to show improvement in unselected groups of patients, but can and do demonstrate efficacy in specific subgroups. Identifying clinical phenotypes based on statistical methods such as cluster analysis and factor analysis will allow us to reclassify and improve our understanding of the clinical presentations of asthma and COPD, well beyond the traditional subjective categories of intrinsic or extrinsic asthma, or pink puffers and blue bloaters. This approach to classification can then be related to specific molecular mechanisms suggested by analysis of biospecimens, and functional abnormalities assessed by physiological changes. An example of this approach was recently taken using bronchoalveolar lavage specimens from patients with asthma and found that panels of cytokines could be grouped by cluster analysis into four molecular phenotypes, one of which strongly corresponded to patients with severe asthma (11). Using systems methods, continued iteration will be necessary to further refine phenotypes based on molecular, cellular, and physiological profiles that relate to significant clinical outcomes. This strategy has been successful in some areas, such as brain cancer (29), where genetic profiling has improved classification of gliomas based on molecular subtypes that correspond to survival better than histological classification.

This systems biology approach will undoubtedly also generate novel hypotheses that can then be further tested in clinical trials. By isolating more clinically relevant phenotypes based on underlying disease mechanism, the systems approach may radically change the way in which these clinical trials are performed. In order to deal with inherent patient heterogeneity, phase III drug trials currently must be highly selective in their inclusion criteria, and yet still involve large numbers of patients treated and monitored over long periods of time. Thus many patients may be excluded from study, and these trials are costly and time consuming. A better strategy is to keep initial clinical trials large and simple, allowing broad inclusion of subjects and, through careful phenotyping, isolation of distinct patient subtypes that respond to the intervention being tested. Subsequent, more definitive trials can then focus on specific patient phenotypes and monitor changes in relevant biomarkers and other outcomes via fast, high-throughput methods, allow-
ing the sample size and length of such trials to be reduced, and bringing to practice sooner new therapeutic strategies (54).

CONCLUSIONS

Applying a systems biology approach to lung disease makes sense, because the lung is a highly complex biological organ, which will easily evade understanding by using only a traditional reductionist research approach. This has been revealed repeatedly by the many failed clinical attempts at treating asthma, COPD, acute lung injury, and other diseases by targeting only one molecule or mediator. However, application of systems biology to lung disease is truly an “emerging” field. Progress has been made in acquiring high-throughput data on molecular and genetic profiles of disease, in measuring and developing analytic methods to understand functional dynamics, and in applying unbiased statistical methods to discover and redefine clinical phenotypes of disease. The challenge remains in integrating this information into a comprehensive model that can be tested and refined, and ultimately related to clinical outcomes. This process will lead to better diagnosis and individualized treatment. In fact, the promise of systems biology is a new approach to medicine that has been described as “predictive, personalized, preventive and participatory” (65). It will be predictive in the sense of allowing better prognostic information based on a more individual understanding of disease; personalized because individual phenotypes will be better characterized and related to treatment outcomes; preventive because of the predictive nature and subsequent better understanding of mechanism and consequent intervention to prevent disease; and participatory because in the true sense of a complex network, individual behavior, stress and other lifestyle factors, as well as more global public health issues, will also play a role in disease pathogenesis (7). Of course, as with any new paradigm, barriers to successful application of the systems approach must be addressed. These include cost-effective methods to gather large bodies of data, and careful techniques to analyze these data in order to avoid false discovery that may contribute to confusion, not understanding (3, 12). Integrating systems biology into medical research and practice will require sufficient funding, infrastructure, academic recognition, and facilitation of the transdisciplinary approach required (15). A major issue will be one of effective communication among scientists, clinicians, and policy makers across many disciplines, both in academia and in industry (7). Common, open-access databases, analytic tools, and tailored knowledge management systems, including the language of systems biology with pathway and network analyses, are already in use and under further development in an effort to facilitate the growth and development of the systems biology approach in medicine (7).

ACKNOWLEDGMENTS

The authors thank Gary J. Nelson for creating the illustration in Fig. 1.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES


