JCB*i*

JOURNAL OF
CLINICAL BIOINFORMATICS

REVIEW   **Open Access**

Significance of bioinformatics in research of chronic obstructive pulmonary disease

Hong Chen[1] and Xiangdong Wang[1,2]*

Abstract

Chronic obstructive pulmonary disease (COPD) is an inflammatory disease characterized by the progressive deterioration of pulmonary function and increasing airway obstruction, with high morality all over the world. The advent of high-throughput omics techniques provided an opportunity to gain insights into disease pathogenesis and process which contribute to the heterogeneity, and find target-specific and disease-specific therapies. As an interdispline, bioinformatics supplied vital information on integrative understanding of COPD. This review focused on application of bioinformatics in COPD study, including biomarkers searching and systems biology. We also presented the requirements and challenges in implementing bioinformatics to COPD research and interpreted these results as clinical physicians.

Keywords: bioinformatics, genomics, proteomics, chronic obstructive pulmonary disease, biomarkers

Introduction

Chronic obstructive pulmonary disease (COPD) is an inflammatory disease characterized by the progressive deterioration of pulmonary function and increasing airway obstruction [1,2]. It can be caused by inflammatory responses triggered by noxious particles or gas, most commonly from tobacco smoking and is accompanied by chronic bronchitis and emphysema [3,4]. Some patients go on to require long-term oxygen therapy or even lung transplantation [3]. COPD was ranked as fourth leading cause of death worldwide and is estimated to become the top third cause of mortality by 2020 [5]. According to the data in China, COPD ranks as the fourth leading cause of death in urban areas and third in rural areas[6]. The high mortality and morbidity with COPD, and its chronic progressive nature, have promoted the need to investigate the underlying mechanisms and identify biomarkers for diagnosis, prognosis and drug target.

The understanding of COPD increased by advanced molecular biology approaches, genetically modified animals, virally administered genes, and high-throughput transcriptional profiling approaches. High-throughput

methodologies, such as genomics and proteomics, are commonly used. The variety of data from biology, mainly in the form of DNA, RNA and protein sequences is putting heavy demand in computer sciences and computational biology. Bioinformatics, including many sub-disciplines, such as genomics, proteomics and system biology, is an integration of mathematical, statistical, and computational methods to analyze biological, biochemical, and biophysical data. Compared to wet-lab method, bioinformatics focused on data mining via computational means. Sophisticated bioinformatics techniques are developed to analyze the vast amount of data generated from genomics and proteomics studies, such as gene and protein function, interactions and metabolic and regulatory pathways. However, there is still a great challenge to combine the computer figures with clinical data for both bench-scientists and bedside-physicians.

In COPD studies, there are usually three ways to analyze 'omics' data: 1) search correlation between single gene or protein and some clinical features in order to find diagnostic or prognostic biomarkers; 2) integrate clinical and wet-lab information, or omics data from different levels for database establishment and computational models. In this current review, we discussed application of bioinformatics in COPD study. We also presented the requirements and challenges in implementing bioinformatics to COPD research, and gave

---

* Correspondence: xiangdong.wang@telia.com
[1]Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University, Shanghai, China
Full list of author information is available at the end of the article

BioMed Central

some suggestions on how to interpret these results as clinical physicians.

Application of bioinformatics in biomarker searching

The diagnosis of COPD is based on the presence of typical symptoms of cough and shortness of breath, together with the presence of risk factors, and is confirmed by spirometry. Therefore, searching for better biomarkers with high specificity and sensitivity indicating the staging and severity of COPD remain as major concerns for clinical physicians. The main value of biomarkers in COPD would be in early diagnosis and to provide the early proof of drug efficacy during the treatment [7]. As a biomarker for COPD, it is expected to be detected in human lung fluids or tissues, sensitive to the progress of COPD, disease-specific to COPD and associated with the status of patients [8]. In these researches, selected genes or proteins usually combined with several clinical features, such as disease susceptibility, lung function, via statistical methods, i.e. logistic regression.

Genomics

It is believed that many genetic factors increase a person's risk of developing COPD[9]. The high mortality and morbidity associated with COPD, and its chronic and progressive nature, has prompted the use of molecular genetic studies in an attempt to identify susceptibility factors for the disease. The advent of high-throughput methodologies to study genetic background variability, epigenetic regulation allowed us to explain the individual variability in the susceptibility of human diseases.

Single nucleotide polymorphism (SNP) was a common method in COPD study. These analyses usually study a group of candidate genes, and then perform statistical test in different populations. The gene-related susceptibility can be approached by testing or unbiased study designs [10]. By SNP microarray, many genetic factors were found to be related with the individual risk of developing COPD [9]. Apart from recognized deficiency of alpha1antitrypsin [9], genomics in COPD found that other gene alleles, such as IREB2[11], CYP2E1 and NAT2[12], CYP1A1, CYP1A2 and CYBA[13], TNF-α [14] were associated with COPD susceptibility. PIM3 allele of the alpha1antitrypsin gene had an association with the pathogenesis of COPD in the Indian population [15]. The polymorphisms in SP-A1 and SP-A2[16], COX2 and p53 risk-alleles [17]might be genetic factors contributing to the susceptibility to COPD. On the other hand, COPD could influence single gene expression as well, such as cathepsin inhibitory cystatin A[18].

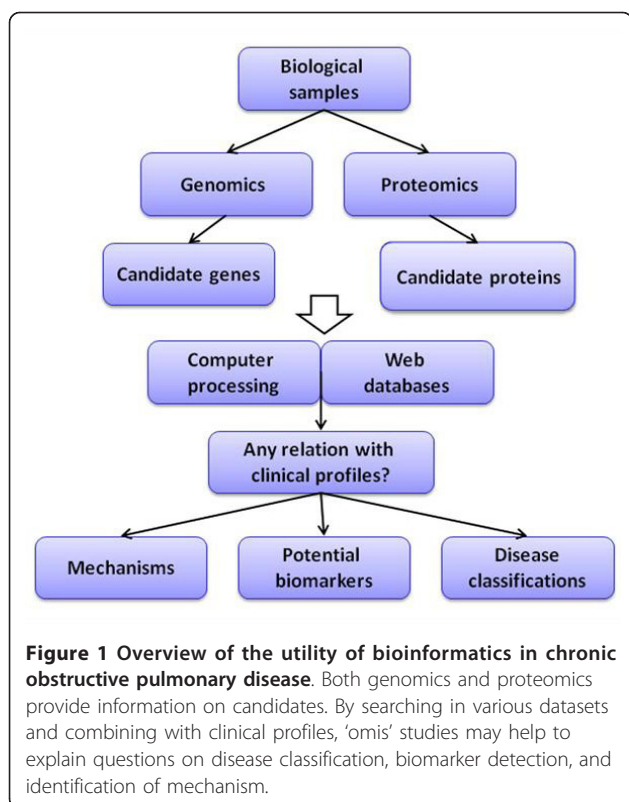COPD is featured by decline in lung function in disease progression. Therefore, except susceptibility, other case-control genomics studies focused on the association of several gene alleles with lung function. 105V/114V alleles of GSTP1 and 113H/139H alleles of mEPHX and the combination of genotypes with same alleles were associated with imbalanced oxidative stress and lung function in patients [19]. Polymorphisms in ADAM33 were associated with COPD and lung function decline in long-term smokers [20] and general population [21]. The variants and their combinations of eNOS -786C, -922G, and 4A alleles in endothelial cells contribute to disturbed pulmonary function and oxidative stress in COPD[22].

An interleukin13 polymorphism in the promoter region may modulate the adverse effects of cigarette smoking on pulmonary function in long-term cigarette smokers [23]. These genomics findings suggested that environmental influences were important in COPD. Genetic polymorphisms contributed to the development of COPD, especially to the declined lung function. The diversity in human genes could help us to understand the susceptibility among different ethics and different populations. The dissection of the genetic basis of complex diseases and the development of highly individualized therapies remain lofty but achievable goals [24].

Proteomics

Proteomics is the systematic study of the many and diverse properties of protein profiles in a parallel manner with the aim of providing detailed descriptions of the structure, function and control of biological systems in health and disease [25]. A major research objective is to search for biomarkers in complex biological fluids. The proteomic analysis highlights the avenues to investigate protein profiles of cells, biopsies and fluids, explore protein-based mechanisms of human diseases, define subgroups of disease, and identify novel biomarkers for diagnosis, therapy and prognosis of multiple diseases and discover new targets for drug development. In particular, the application of complementary approaches, including gel- and liquid chromatography mass spectrometry-based proteomic techniques on sputum and/or bronchoalveolar lavage may provide a better understanding of the proteome differentially expressed among the courses, severities and populations of COPD [7]. We have previously reviewed the clinical studies on COPD proteomics, highlighted the proteomic-oriented methods applied and evaluated the diagnostic or prognostic values of potential biomarkers [8]. Those studies mainly focused on disease classification, biomarker detection, or identification of mechanism, while the three components are related with each other in COPD (Figure 1).

An important goal of proteomic studies is to understand biological roles of specific proteins and develop

Figure 1 Overview of the utility of bioinformatics in chronic obstructive pulmonary disease. Both genomics and proteomics provide information on candidates. By searching in various datasets and combining with clinical profiles, 'omis' studies may help to explain questions on disease classification, biomarker detection, and identification of mechanism.

new therapeutic targets [26]. Although there are few proteomics studies performed in COPD patients, several potential proteins have been regarded as biomarkers. For example, matrix metalloproteinase (MMP)-13 and thioredoxin-like 2 in lungs increased in patients with COPD [27]. The serine and MMP proteinase network was considered as an important feature in predicting clinical worsening of airway obstruction [28]. Pulmonary surfactant A was found to link to the pathogenesis of COPD and could be considered as a potential COPD biomarker [29]. Proteomic screening of sputum yields potential biomarkers of inflammation [30]. Airway and parenchymal phenotypes of COPD were suggested to be associated with unique systemic serum biomarker profiles [31]. The utility of proteomic profiling would improve the understanding of molecular mechanisms involved in cigarette smoking-related COPD by identifying plasma proteins that correlate with declined lung function [32]. The concentrations of neutrophil defensins 1 and 2, calgranulin A, and calgranulin B were elevated in smokers with COPD when compared to asymptomatic smokers [33]. Other candidates, like serum amyloid A [34], plasma retinal-binding protein, apolipoprotein E, inter-alpha-trypsininhibitor heavy chain H4, and glutathione peroxidase[35], were also been detected in plasma in COPD patients by proteomics approaches.
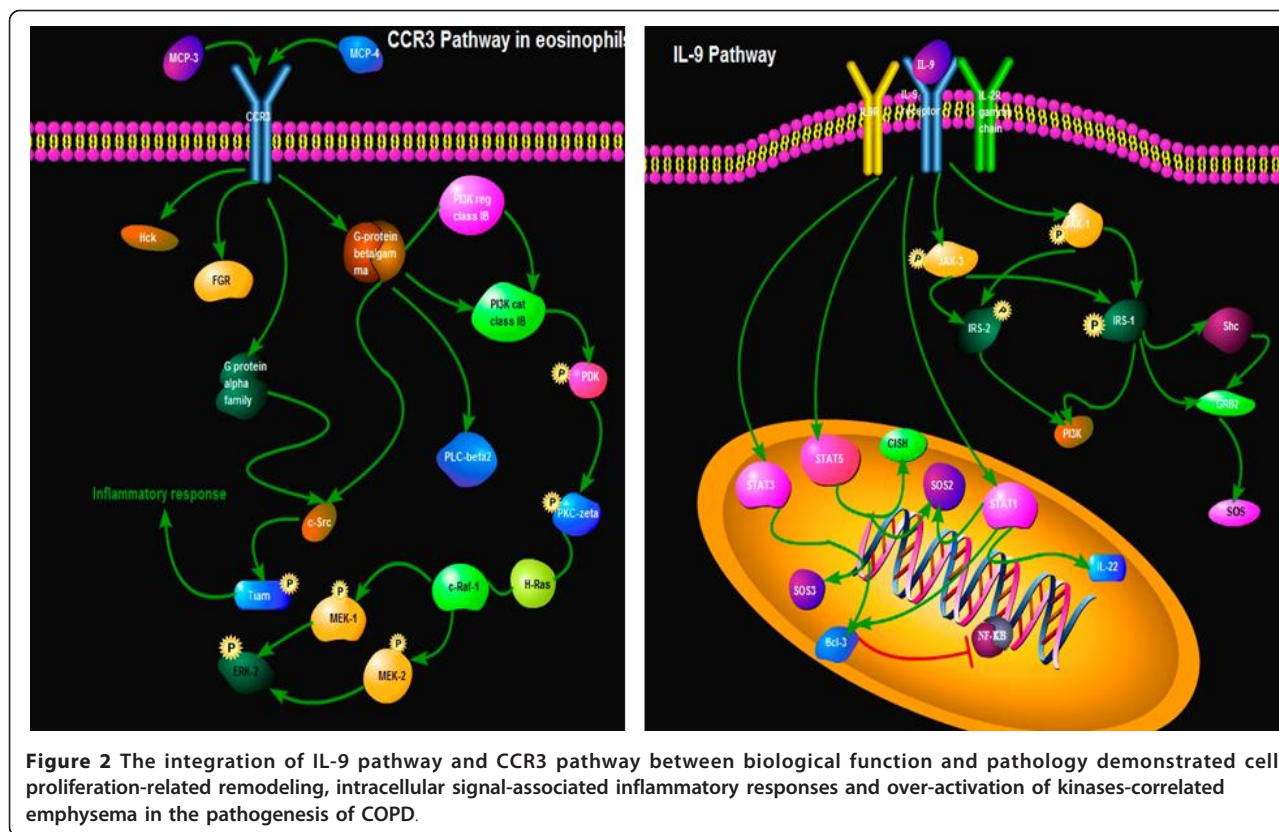
Metabolomics

Metabolomics is a global way to understanding regulation of metabolic pathways and metabolic networks of a biologic system[36]. Metabolites trigonelline, hippurate and formate in urinary were identified to be associated with baseline lung function of COPD patients and considered to reflect lifestyle differences affecting overall health [37]. Another way to analyze omics data is clustering data to form a specific pattern for different groups by principal component analysis (PCA). Combination of PCA and metabolomics identified the metabolic fingerprint of exhaled breath condensate of COPD patients [38].

Multiplexed ELISA

A recent advancement in ELISA is the multiplexed ELISA which could determine multiple proteins within a single tissue sample. It has recently been shown to be more sensitive than standard ELISA once optimized for a particular cytokine [39-41] and could be a promising diagnostic assay in lung diseases [42]. Moreover, the measurement of multiple cytokines is required for many diseases, particularly those like COPD that arise from a complex process of initiation and progression of inflammation network. Simultaneous detection of multiple cytokines will undoubtedly provide a more powerful tool to quantifiably measure cytokines in different stages of COPD. With help of this high-throughput measurement platform, we could integrate both biologic and clinical data to inform predictive multiscale models ranging from the molecular to the organ levels, as shown in Figure 2. For example, the integration of IL-9 pathway and CCR3 pathway between biological function and pathology demonstrated cell proliferation-related remodeling, intracellular signal-associated inflammatory responses and over-activation of kinases-correlated emphysema in the pathogenesis of COPD. We propose this new way will increase our insight into disease process and have great potential to identify new biomarkers for disease diagnosis as well as novel therapeutic targets.

Systems biology and database establishment

The difficulties encountered whilst exploring pathogenesis and searching for biomarkers may be due in part to the complex nature of COPD, which comprises a broad spectrum of histopathological findings and respiratory symptoms [43]. All genetic information and molecular knowledge need to be semantically incorporated and associated with clinical and experimental data. System biology in COPD (as reviewed before[44]) presented a manifold understanding of the complexity of COPD, therefore advanced biomedical research and drug development. This approach relies on global genome, transcriptome, proteome, and metabolome data sets

Figure 2 The integration of IL-9 pathway and CCR3 pathway between biological function and pathology demonstrated cell proliferation-related remodeling, intracellular signal-associated inflammatory responses and over-activation of kinases-correlated emphysema in the pathogenesis of COPD.

collected in cross-sectional patient cohorts with high-throughput measurement platforms and integrated with biologic and clinical data to inform predictive multidimensional models ranging from the molecular to the organ levels[45].

Comandini etc. [46] assayed a number of published studies by creating a smoker datasets on which to perform data-mining analysis. They utilized Ingenuity Pathways Analysis, a web-based application that enables identifying relationships, biological mechanisms, functions, and pathways of relevance associated with the molecules under study. Their findings supported the central role of anti-oxidant genes in smoking population and suggested Nrf2 may be a COPD risk biomarker. A brand new knowledge base was generated from clinical and experimental data for COPD based on BioXM software platform [47]. This integrated database reduced implementation time and effort for the knowledge base compared to similar systems and provided a free, comprehensive, easy to use resource for all COPD related clinical research.

Clinical profiles could also be considered as a form of omics data since it provided a large quantity of patients' information in a direct conservative way. An in-silico research applied various explorative analysis techniques (PCA, MCA, MDS) and unsupervised clustering

methods (KHM) to study a large dataset, acquired from 415 COPD patients, to assess the presence of hidden structures in data corresponding to the different COPD phenotypes observed in clinical practice. This study may be considered as a methodological example showing possible applications of intelligent data analysis and visual exploratory techniques to investigate clinical aspects of chronic pathologies where a mathematical referring model is generally missing[48].
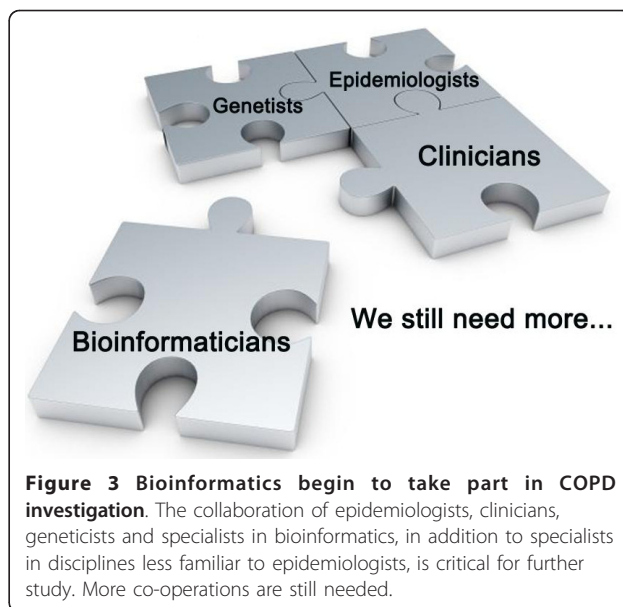
How to understand bioinformatics as clinical physicians

COPD has been approached by genomic and proteomic technologies to allow us to identify patterns of gene/protein expression that track with clinical disease or to identify new pathways involved in disease pathogenesis. The results from these initial studies highlight the potential for these omics approaches to reveal novel insights into the pathogenesis of COPD and provide new tools to improve diagnosis, clinical classification, course prediction, and response to therapy. Existing knowledge such as genotype-phenotype relations or signal transduction pathways must be semantically integrated and dynamically organized into structured networks that are connected with clinical and experimental data. This will require collaboration among

multidisciplinary groups with expertise in the respective technologies, bioinformatics, and clinical medicine for the disease. More and more clinical physicians began to realize the promise of these studies and the potential to revolutionize the diagnosis and treatment of COPD, while obstacles still existed between these laboratory findings and their applications in clinical practice. Omics results need to be interpreted by translational medicine and systems biology. Since the vocabulary of systems biology is different than that of molecular biology or clinical research, the biggest challenge is the shift in thinking[49].

Instead of a single-factor approach, which is highly effective in the lab, we need to think globally as clinical physicians. We need to shift from an approach that tries to explain lung pathogenesis by "one molecule, one cell type" to approach that looks at the network of interactions between multiple molecules, pathways, and cells. Given that all the samples were collected from human, it would be of great significance to standardize patient groups. Criteria of clinical informatics and medical informatics, including age, gender, smoking history, staging, complications and clinical signs as well as examinations, should be fully considered before and after any omics investigation. We also need to pay attention to the relation between clinical data and laboratory findings. For this to occur, a well-done history and physical examination would be helpful to supplement these laboratory figures by providing multiple features of human COPD.
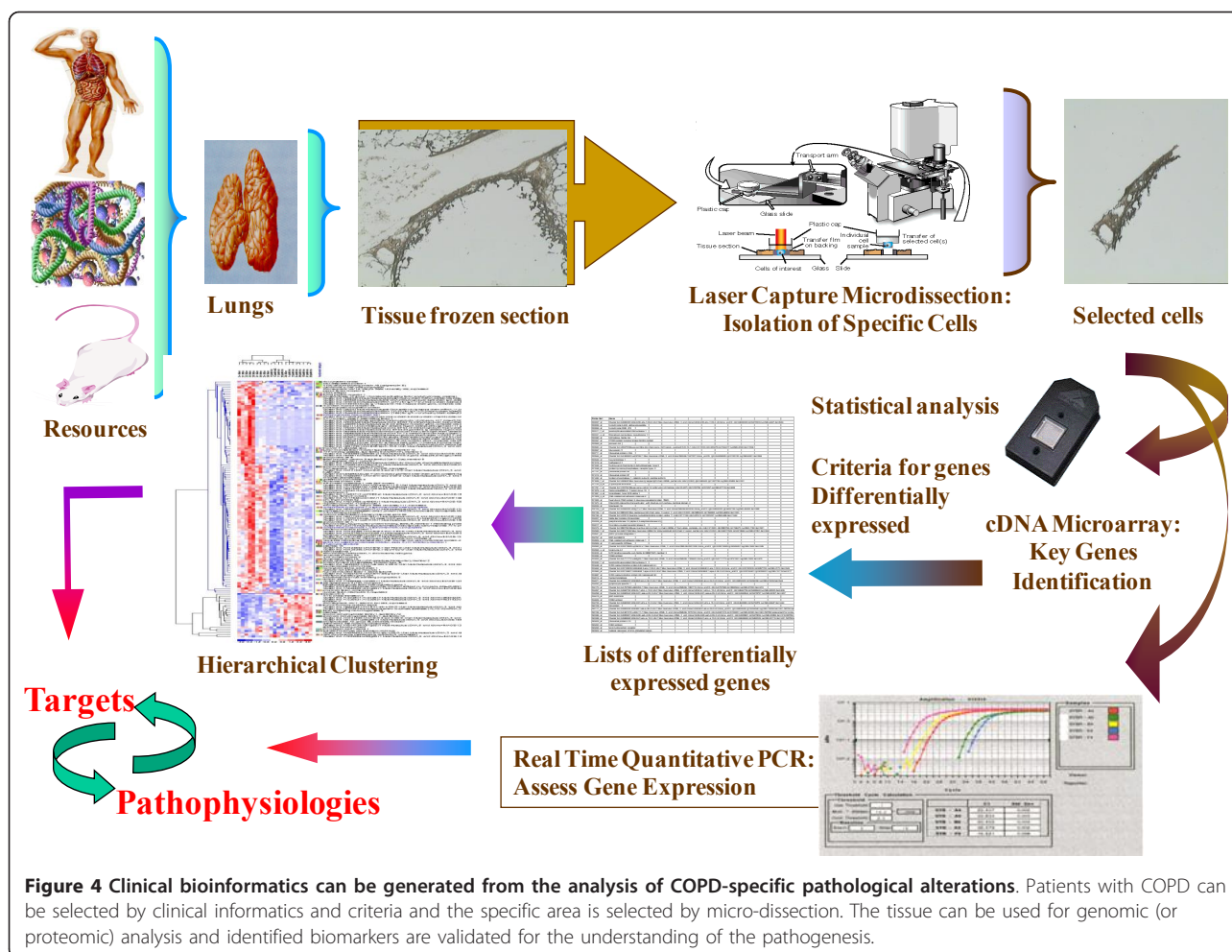
Although all of these exciting technological advances that exponentially increase the levels of knowledge about every disease and model serve as facilitators of integration, they do not inherently provide integrative models of disease. Therefore, we proposed that digitalizing essential clinical profiles, such as symptoms and signs, by questionnaires and/or scores, would provide direct vision for physicians and shrink the distance between lab discovery and clinical condition. The combination of epidemiologists, clinicians, geneticists and specialists in bioinformatics, in addition to specialists in disciplines less familiar to epidemiologists, is critical to be prepared for new phenotypic characterizations based on transcriptome and proteome [50]. Even though we have spotted considerable advancement in bioinformatics, it still calls for more collaboration to fulfill its potential (Figure 3). Interdisciplinary teams should allow us to access omics datasets integratively and generate a global model of COPD. It is important to have a special attention from proteomic scientists to explore the combination between advanced proteomic biotechnology, clinical proteomics, tissue imaging and profiling, and organ dysfunction score systems, to improve the clinical outcomes of these patients [51]. There is still a great

Figure 3 Bioinformatics begin to take part in COPD investigation. The collaboration of epidemiologists, clinicians, geneticists and specialists in bioinformatics, in addition to specialists in disciplines less familiar to epidemiologists, is critical for further study. More co-operations are still needed.

need to explore the COPD-specific and/or related transcriptional factors and regulation networks generated from omics and bioinformatics like in other diseases [41].

Conclusions

The use of high-throughput techniques for gene and protein expression profiling and of computerized databases has become a mainstay of biomedical research. There is a need to perform omics studies on patients with COPD, describing the association with the disease in terms of specificity, severity, progress and prognosis and monitoring the efficacy of therapies. These omics analysis highlight the ways to investigate protein profiles of cells, biopsies and fluids, explore protein-based mechanisms of human diseases, identify novel biomarkers for diagnosis, therapy and prognosis of multiple diseases and discover new targets for drug development, as shown in Figure 3. Although the number of clinical studies on COPD is limited, they still serve as the outstanding initiation for proteomic research in such a complex disease. The analysis of protein profiles that are up- or down-regulated, modified, secreted in the airways during the disease may yield vital evidences to understand the pathogenesis and discover new therapeutic targets for the disease (Figure 4) [52]. With many guidelines now in place and model studies on which to design future experiments, there is reason to be optimistic that candidate protein biomarkers will be discovered using proteomics and translated into clinical assays[53]. With better study design standardization and the implementation of novel technologies to reach the optimal research standard, there is enough reason be optimistic

Figure 4 Clinical bioinformatics can be generated from the analysis of COPD-specific pathological alterations. Patients with COPD can be selected by clinical informatics and criteria and the specific area is selected by micro-dissection. The tissue can be used for genomic (or proteomic) analysis and identified biomarkers are validated for the understanding of the pathogenesis.

about the future of omics research and its clinical implications [54]. Clinical bioinformatics on COPD could be achieved from the combination of clinical informatics, medical informatics, bioinformatics and informatics by collaborations among clinicians, bioinformaticians, computer scientists, biologists, and mathematicians.

Author details
[1]Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University, Shanghai, China. [2]Biomedical Research Center, Zhongshan Hospital, Fudan University, Shanghai, China.

Authors' contributions
HC mainly wrote the manuscript, as well as the revision. XDW conceived of the review and edited and wrote part of the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
1. Celli BR, MacNee W: Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004, **23(6)**:932-946.
2. Pauwels RA, Rabe KF: Burden and clinical features of chronic obstructive pulmonary disease (COPD). *Lancet* 2004, **364(9434)**:613-620.
3. Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, Cherniack RM, Rogers RM, Sciurba FC, Coxson HO, *et al*: The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004, **350(26)**:2645-2653.
4. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, *et al*: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007, **176(6)**:532-555.
5. Lopez AD, Murray CC: The global burden of disease, 1990-2020. *Nat Med* 1998, **4(11)**:1241-1243.
6. Fang X, Wang X, Bai C: COPD in China: the burden and importance of proper management. *Chest* 2011, **139(4)**:920-929.
7. Casado B, Iadarola P, Luisetti M, Kussmann M: Proteomics-based diagnosis of chronic obstructive pulmonary disease: the hunt for new markers. *Expert Rev Proteomics* 2008, **5(5)**:693-704.
8. Chen H, Wang D, Bai C, Wang X: Proteomics-based biomarkers in chronic obstructive pulmonary disease. *J Proteome Res* 2010, **9(6)**:2798-2808.
9. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS: Global strategy for the diagnosis, management, and prevention of chronic obstructive

pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001, **163**(5):1256-1276.

10. Silverman EK, Spira A, Pare PD: **Genetics and genomics of chronic obstructive pulmonary disease.** *Proc Am Thorac Soc* 2009, **6**(6):539-542.

11. DeMeo DL, Mariani T, Bhattacharya S, Srisuma S, Lange C, Litonjua A, Bueno R, Pillai SG, Lomas DA, Sparrow D, et al: **Integration of genomic and genetic approaches implicates IREB2 as a COPD susceptibility gene.** *Am J Hum Genet* 2009, **85**(4):493-502.

12. Arif E, Vibhuti A, Alam P, Deepak D, Singh B, Athar M, Pasha MA: **Association of CYP2E1 and NAT2 gene polymorphisms with chronic obstructive pulmonary disease.** *Clin Chim Acta* 2007, **382**(1-2):37-42.

13. Vibhuti A, Arif E, Mishra A, Deepak D, Singh B, Rahman I, Mohammad G, Pasha MA: **CYP1A1, CYP1A2 and CYBA gene polymorphisms associated with oxidative stress in COPD.** *Clin Chim Acta* 2010, **411**(7-8):474-480.

14. Zhan P, Wang J, Wei SZ, Qian Q, Qiu LX, Yu LK, Song Y: **TNF-308 gene polymorphism is associated with COPD risk among Asians: meta-analysis of data for 6,118 subjects.** *Mol Biol Rep* 2011, **38**(1):219-227.

15. Gupta J, Bhadoria DP, Lal MK, Kukreti R, Chattopadhaya D, Gupta VK, Dabur R, Yadav V, Chhillar AK, Sharma GL: **Association of the PIM3 allele of the alpha-1-antitrypsin gene with chronic obstructive pulmonary disease.** *Clin Biochem* 2005, **38**(5):489-491.

16. Saxena S, Kumar R, Madan T, Gupta V, Muralidhar K, Sarma PU: **Association of polymorphisms in pulmonary surfactant protein A1 and A2 genes with high-altitude pulmonary edema.** *Chest* 2005, **128**(3):1611-1619.

17. Arif E, Vibhuti A, Deepak D, Singh B, Siddiqui MS, Pasha MA: **COX2 and p53 risk-alleles coexist in COPD.** *Clin Chim Acta* 2008, **397**(1-2):48-50.

18. Butler MW, Fukui T, Salit J, Shaykhiev R, Mezey JG, Hackett NR, Crystal RG: **Modulation of Cystatin A Expression in Human Airway Epithelium Related to Genotype, Smoking, COPD, and Lung Cancer.** *Cancer Res* 2011, **71**(7):2572-2581.

19. Vibhuti A, Arif E, Deepak D, Singh B, Qadar Pasha MA: **Genetic polymorphisms of GSTP1 and mEPHX correlate with oxidative stress markers and lung function in COPD.** *Biochem Biophys Res Commun* 2007, **359**(1):136-142.

20. Sadeghnejad A, Ohar JA, Zheng SL, Sterling DA, Hawkins GA, Meyers DA, Bleecker ER: **Adam33 polymorphisms are associated with COPD and lung function in long-term tobacco smokers.** *Respir Res* 2009, **10**:21.

21. van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Schouten JP, Boezen HM: **A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population.** *Am J Respir Crit Care Med* 2005, **172**(3):329-333.

22. Arif E, Ahsan A, Vibhuti A, Rajput C, Deepak D, Athar M, Singh B, Pasha MA: **Endothelial nitric oxide synthase gene variants contribute to oxidative stress in COPD.** *Biochem Biophys Res Commun* 2007, **361**(1):182-188.

23. Sadeghnejad A, Meyers DA, Bottai M, Sterling DA, Bleecker ER, Ohar JA: **IL13 promoter polymorphism 1112C/T modulates the adverse effect of tobacco smoking on lung function.** *Am J Respir Crit Care Med* 2007, **176**(8):748-752.

24. Desai AA, Hysi P, Garcia JG: **Integrating genomic and clinical medicine: searching for susceptibility genes in complex lung diseases.** *Transl Res* 2008, **151**(4):181-193.

25. Patterson SD, Aebersold RH: **Proteomics: the first decade and beyond.** *Nat Genet* 2003, **33**(Suppl):311-323.

26. Marko-Varga G, Fehniger TE: **Proteomics and disease–the challenges for technology and discovery.** *J Proteome Res* 2004, **3**(2):167-178.

27. Lee EJ, In KH, Kim JH, Lee SY, Shin C, Shim JJ, Kang KH, Yoo SH, Kim CH, Kim HK, et al: **Proteomic analysis in lung tissue of smokers and COPD patients.** *Chest* 2009, **135**(2):344-352.

28. Sepper R, Prikk K: **Proteomics: is it an approach to understand the progression of chronic lung disorders?** *J Proteome Res* 2004, **3**(2):277-281.

29. Ohlmeier S, Vuolanto M, Toljamo T, Vuopala K, Salmenkivi K, Myllarniemi M, Kinnula VL: **Proteomics of human lung tissue identifies surfactant protein A as a marker of chronic obstructive pulmonary disease.** *J Proteome Res* 2008, **7**(12):5125-5132.

30. Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC, Innes JA, Porteous DJ, Greening AP: **Sputum proteomics in inflammatory and suppurative respiratory diseases.** *Am J Respir Crit Care Med* 2008, **178**(5):444-452.

31. Bon JM, Leader JK, Weissfeld JL, Coxson HO, Zheng B, Branch RA, Kondragunta V, Lee JS, Zhang Y, Choi AM, et al: **The influence of**

radiographic phenotype and smoking status on peripheral blood biomarker patterns in chronic obstructive pulmonary disease. *PLoS One* 2009, **4**(8):e6865.

32. York TP, van den Oord EJ, Langston TB, Edmiston JS, McKinney W, Webb BT, Murrelle EL, Zedler BK, Flora JW: **High-resolution mass spectrometry proteomics for the identification of candidate plasma protein biomarkers for chronic obstructive pulmonary disease.** *Biomarkers* 2010, **15**(4):367-377.

33. Merkel D, Rist W, Seither P, Weith A, Lenter MC: **Proteomic study of human bronchoalveolar lavage fluids from smokers with chronic obstructive pulmonary disease by combining surface-enhanced laser desorption/ionization-mass spectrometry profiling with mass spectrometric protein identification.** *Proteomics* 2005, **5**(11):2972-2980.

34. Bozinovski S, Hutchinson A, Thompson M, Macgregor L, Black J, Giannakis E, Karlsson AS, Silvestrini R, Smallwood D, Vlahos R, et al: **Serum amyloid a is a biomarker of acute exacerbations of chronic obstructive pulmonary disease.** *Am J Respir Crit Care Med* 2008, **177**(3):269-278.

35. Bandow JE, Baker JD, Berth M, Painter C, Sepulveda OJ, Clark KA, Kilty I, VanBogelen RA: **Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies–COPD biomarker discovery study.** *Proteomics* 2008, **8**(15):3030-3041.

36. Nicholson JK, Connelly J, Lindon JC, Holmes E: **Metabonomics: a platform for studying drug toxicity and gene function.** *Nat Rev Drug Discov* 2002, **1**(2):153-161.

37. McClay JL, Adkins DE, Isern NG, O'Connell TM, Wooten JB, Zedler BK, Dasika MS, Webb BT, Webb-Robertson BJ, Pounds JG, et al: **(1)H nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function.** *J Proteome Res* 2010, **9**(6):3083-3090.

38. de Laurentiis G, Paris D, Melck D, Maniscalco M, Marsico S, Corso G, Motta A, Sofia M: **Metabonomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy.** *Eur Respir J* 2008, **32**(5):1175-1183.

39. Trune DR, Larrain BE, Hausman FA, Kempton JB, Macarthur CJ: **Simultaneous measurement of multiple ear proteins with multiplex ELISA assays.** *Hear Res* 2010.

40. Elshal MF, McCoy JP: **Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA.** *Methods* 2006, **38**(4):317-323.

41. Ray CA, Bowsher RR, Smith WC, Devanarayan V, Willey MB, Brandt JT, Dean RA: **Development, validation, and implementation of a multiplex immunoassay for the simultaneous determination of five cytokines in human serum.** *J Pharm Biomed Anal* 2005, **36**(5):1037-1044.

42. Farlow EC, Patel K, Basu S, Lee BS, Kim AW, Coon JS, Faber LP, Bonomi P, Liptay MJ, Borgia JA: **Development of a multiplexed tumor-associated autoantibody-based blood test for the detection of non-small cell lung cancer.** *Clin Cancer Res* 2010, **16**(13):3452-3462.

43. Bhattacharya S, Srisuma S, Demeo DL, Shapiro SD, Bueno R, Silverman EK, Reilly JJ, Mariani TJ: **Molecular biomarkers for quantitative and discrete COPD phenotypes.** *Am J Respir Cell Mol Biol* 2009, **40**(3):359-367.

44. 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**(9):1129-1137.

45. Auffray C, Adcock IM, Chung KF, Djukanovic R, Pison C, Sterk PJ: **An integrative systems biology approach to understanding pulmonary diseases.** *Chest* 2010, **137**(6):1410-1416.

46. Comandini A, Marzano V, Curradi G, Federici G, Urbani A, Saltini C: **Markers of anti-oxidant response in tobacco smoke exposed subjects: a data-mining review.** *Pulm Pharmacol Ther* 2010, **23**(6):482-492.

47. Maier D, Kalus W, Wolff M, Kalko SG, Roca J, Marin de Mas I, Turan N, Cascante M, Falciani F, Hernandez et al: **Knowledge management for systems biology a general and visually driven framework applied to translational medicine.** *BMC Syst Biol* 2011, **5**:38.

48. Paoletti M, Camiciottoli G, Meoni E, Bigazzi F, Cestelli L, Pistolesi M, Marchesi C: **Explorative data analysis techniques and unsupervised clustering methods to support clinical assessment of Chronic Obstructive Pulmonary Disease (COPD) phenotypes.** *J Biomed Inform* 2009, **42**(6):1013-1021.

49. Studer SM, Kaminski N: **Towards systems biology of human pulmonary fibrosis.** *Proc Am Thorac Soc* 2007, **4**(1):85-91.

50. Kauffmann F: **Post-genome respiratory epidemiology: a multidisciplinary challenge.** *Eur Respir J* 2004, **24**(3):471-480.

51. Wang X, Adler KB, Chaudry IH, Ward PA: **Better understanding of organ dysfunction requires proteomic involvement.** *J Proteome Res* 2006, **5(5)**:1060-1062.

52. Plymoth A, Yang Z, Lofdahl CG, Ekberg-Jansson A, Dahlback M, Fehniger TE, Marko-Varga G, Hancock WS: **Rapid proteome analysis of bronchoalveolar lavage samples of lifelong smokers and never-smokers by micro-scale liquid chromatography and mass spectrometry.** *Clin Chem* 2006, **52(4)**:671-679.

53. Reisdorph NA, Reisdorph R, Bowler R, Broccardo C: **Proteomics methods and applications for the practicing clinician.** *Ann Allergy Asthma Immunol* 2009, **102(6)**:523-529.

54. Lin JL, Bonnichsen MH, Nogeh EU, Raftery MJ, Thomas PS: **Proteomics in detection and monitoring of asthma and smoking-related lung diseases.** *Expert Rev Proteomics* 2010, **7(3)**:361-372.