Chronic Obstructive Pulmonary Disease and Proteomics: A Match for Success?
Bruno M. Alexandre and Deborah Penque*
Laboratório de Proteómica, Departamento de Genética, Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA-IP), Lisboa, Portugal

Abstract

Chronic obstructive pulmonary disease (COPD) is characterized by chronic airflow limitation that is not fully reversible even under bronchodilators effect, caused by a mixture of small airway disease and parenchymal destruction. COPD is a major cause of morbidity and mortality in adults, and it is now the fourth leading death cause in the world. Cigarette smoking is the main risk factor for COPD but not all smokers will suffer from COPD, suggesting that genetic and other environmental factors are involved in this pathology.

Current diagnosis is based on spirometry, but there is recurrent debate on fixed spirometric thresholds in use that lead to misdiagnosis and/or classification of COPD. The available treatments are not effective to reduce or suppress the progression of COPD. Hence, there is an urgent need to better understand the molecular mechanisms of COPD pathogenesis to provide clinicians with reliable diagnosis and treatment tools for COPD. Proteomics, defined by the comprehensive study of the proteome, has the potential to respond to this need by providing protein profiles of a particular disease and, at the same time, by identifying specific biomarkers that can be used to better understand, diagnose and manage the disease. Here, we shortly review COPD history and pathology and how proteomics can match COPD for success.

Keywords: Chronic obstructive pulmonary disease; Respiratory disease; Proteomics; Clinical proteomics; Review

Chronic Obstructive Pulmonary Disease

History

Chronic obstructive pulmonary disease (COPD) has certainly long existed, but first reports that may be traced to this disease only dates from the seventeenth century. It was in 1679 that took place the first hypothetical report of COPD cases when Bonet described emphysema as a condition of “voluminous lungs” [1]. Almost a century ahead, Giovanni Morgagni described 19 cases of “turbid” lungs in 1769 and 20 years later an emphysematous lung is illustrated by Matthew Baillie [1]. Early reports of chronic bronchitis were generated in 1814 by Badham who used the word catarrh to refer to the chronic cough and mucus hypersecretion that are key symptoms. He also described bronchiolitis and chronic bronchitis as disabling disorders [1]. The emphysema component of disease was beautifully described by Laënnec (1821) in his Treatise of Diseases of the chest. He recognized that emphysema lungs were hyperinflated and did not empty well [1]. Spirometer was invented in 1846 by John Hutchinson [1]. This device is today absolutely necessary to the correct diagnosis and management of COPD. However, Hutchinson’s instrument only measured vital capacity. A century went by until Tiffeneau was able to add the concept of timed vital capacity as a measure of airflow [2]. Gaensler introduced the concept of the air velocity index based on Tiffeneau’s work and later the forced vital capacity [3], which is the foundation of the FEV1 and FEV1/FVC percent and spirometry became complete as a COPD diagnostic instrument. What was once called chronic obstructive bronchopulmonary disease, chronic airflow obstruction, chronic obstructive lung disease, nonspecific chronic pulmonary disease, and diffuse obstructive pulmonary syndrome, was coined COPD in 1965 by William Briscoe [4].

Definition and prevalence

COPD is characterized by chronic airflow limitation that is not fully reversible even under bronchodilators effect, caused by a mixture of small airway disease - obstructive bronchiolitis - and parenchymal destruction - emphysema. Indeed, main components of COPD are chronic bronchitis and emphysema. Chronic bronchitis is defined by the presence of chronic recurrent increase in bronchial secretions sufficient to cause expectoration. These secretions must be present in most days for a minimum of three months per year for at least two consecutive years and cannot be attributed to other disorders [5]. Noteworthy, not every patient with chronic bronchitis has or will develop chronic airflow limitation [6]. Emphysema is defined anatomically by permanent, destructive enlargement of airspaces distal to the terminal bronchioles without obvious fibrosis [5].

Associated chronic inflammation causes changes and narrowing of the small airways leading to airway remodeling. Parenchyma destruction is responsible for the loss of alveolar attachments and decrease of lung elastic recoil [7]. These changes reduce the ability of the airways to remain open during expiration. No currently available treatments reduce the progression of COPD or suppress the inflammation in small airways and lung parenchyma [8]. According to the Global initiative for chronic Obstructive Lung Disease (GOLD)’s last report, COPD is defined as a preventable and treatable disease with some significant extrapulmonary effects that may contribute to the severity in individual patients [9].

*Corresponding author: Deborah Penque, Ph.D., Laboratório de Proteómica, Departamento de Genética, Edifício INSA II, Instituto Nacional de Saúde Dr. Ricardo Jorge, INSA, I.P., Avenida Padre Cruz, 1649-016 Lisboa, Portugal, Tel: +351 21750 8137; Fax: +351 21752 8410; E-mail: deborah.penque@insula.min-saude.pt

Received December 24, 2011; Accepted February 22, 2012; Published February 27, 2012


Copyright: © 2012 Alexandre BM, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
COPD is a major cause of morbidity and mortality in adults, and its incidence has been increasing worldwide. In 2000, approximately 2.7 million deaths were caused by COPD [10]. At the present time, COPD is the fourth leading cause of death and its prevalence and mortality are expected to continue increasing in next decade [9-11]. Additionally, COPD is the only major cause of death that is increasing in prevalence worldwide [12], while other causes have been declining since 1970 [13]. Not only mortality but also morbidity associated with COPD are often underestimated by healthcare providers and patients as COPD is frequently underdiagnosed and undertreated [14].

In the US, an estimated 12 million individuals, about 4% of the population, were identified with COPD in 2007 [15]. Moreover, another 12 million were estimated to suffer from COPD undiagnosed [15]. In the EU, a 2009 report from the European Federation of Allergy and Airways Diseases Patients Associations revealed that prevalence varies from an estimated 2% (The Netherlands) to more than 10% (Austria and Germany) [16].

**Diagnosis and classification**

The diagnosis of COPD is based on spirometry associated with medical history data, such as smoking, occupational exposure and infection risk factors. The use of postbronchodilator spirometry is recommended to minimize variability [9]. According to the GOLD guidelines, COPD is diagnosed when there is lung obstruction, i.e., when forced expiratory volume in one second to forced vital capacity ratio (FEV1/FVC) is lower than 0.70. COPD severity stage classification is based on FEV1 values (Table 1). However, fixed spirometric thresholds to diagnosis and stage COPD remain under recurrent medical debates [17-20]. Since FEV1/FVC ratio declines with increasing age, a fixed cut-off of 0.70 leads to overdiagnosis of elder patients underdiagnosis among young adults [21,22]. Therefore, new surrogate biomarkers are of great need so that clinicians involved in primary and secondary care are able to attain successful early diagnosis, prognosis and monitoring therapeutic intervention in COPD.

**Risk factors**

Cigarette smoke is the most commonly encountered risk factor for COPD. Cigarette smoking is the leading cause of preventable death worldwide and yet, despite anti-smoking campaign efforts from such organizations as the European Respiratory Society [12], American Thoracic Society [23] or the World Health Organization (WHO) [24], the number of smokers keeps increasing. Thus, global epidemic of tobacco-associated diseases has progressively worsened.

Cigarette smokers have a higher prevalence of respiratory symptoms, lung function abnormalities, a greater rate of decline in forced expiratory volume in the first second, FEV1, and higher death rates for COPD than nonsmokers [9,12]. A 25-year follow up study of the general population concluded that 92% of COPD deaths occurred in subjects who were current smokers at the beginning of the follow up period and that after 25 years of smoking, at least 25% of smokers without initial disease will develop clinically significant and 30-40% will have COPD [25]. The fact that not all smokers develop clinically significant COPD, suggests that genetic factors may modify each individual risk [12].

COPD is a polygenic disease and a classical example of gene-environment interaction [9]. The only proven genetic risk factor for COPD is the hereditary deficiency of α1-antitrypsin, a major circulating inhibitor of serine proteases, in which a smoker will considerably increase the risk for COPD [26]. Gene mutations and polymorphisms have been studied and several candidate genes associated with COPD phenotypes have been reported, but so far none has been validated [27-30]. Occupational dust, outdoor and indoor pollution, socioeconomic status and genetic determinants are also associated with the development of COPD [12].

**Pathology, pathogenesis, pathophysiology and economic burden**

Cigarette smoke and other noxious particles cause amplified lung inflammation in patients that develop COPD. This may induce parenchymal tissue destruction (emphysema) and disturb normal repair and defense mechanisms resulting in small airway inflammation [31,32]. Emphysema and small airway inflammation and damage lead to the enlargement of alveolar air spaces, airway wall fibrosis, loss of elastic recoil, smooth muscle hypertrophy, goblet cell hyperplasia and mucus plugging. Inflammatory exudates accumulate in the small airways lumen due to reduced mucociliary escalator function [31]. The physiological consequences are airflow collapse over expiration leading to airflow obstruction and hyperinflation (air trapping) which ultimately results in characteristic symptom of breathless and progressive airflow limitation that may lead to death. In general, inflammatory and structural changes in the airways increase with disease severity and persist on smoking cessation [33-35].

COPD is characterized by a specific pattern of inflammation which involves neutrophils, macrophages and lymphocytes [7]. These cells

---

**Table 1**: Classification of COPD stages based on spirometry [9]. FEV1 = Forced Expiratory Volume in 1 sec; FVC = Forced Vital Capacity; PaO2 = Arterial Partial Pressure of Oxygen.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Mild COPD</td>
<td>FEV1/FVC &lt; 70%</td>
</tr>
<tr>
<td></td>
<td>FEV1 ≥ 80% predicted</td>
</tr>
<tr>
<td></td>
<td>With or without symptoms</td>
</tr>
<tr>
<td>2: Moderate COPD</td>
<td>FEV1/FVC &lt; 70%</td>
</tr>
<tr>
<td></td>
<td>50% ≤ FEV1 &lt; 80% predicted</td>
</tr>
<tr>
<td></td>
<td>With or without symptoms</td>
</tr>
<tr>
<td>3: Severe COPD</td>
<td>FEV1/FVC &lt; 70%</td>
</tr>
<tr>
<td></td>
<td>30% ≤ FEV1 &lt; 50% predicted</td>
</tr>
<tr>
<td></td>
<td>With or without symptoms</td>
</tr>
<tr>
<td>4: Very severe COPD</td>
<td>FEV1/FVC &lt; 70%</td>
</tr>
<tr>
<td></td>
<td>FEV1 &lt; 30% predicted or &lt; 50% predicted plus presence of chronic respiratory failure (PaO2 &lt; 60 mm Hg while breathing room air at sea level)</td>
</tr>
</tbody>
</table>

**Figure 1**: Pathogenesis of COPD. Source: Barnes PJ. Adapted from www.goldcopd.org.
release inflammatory mediators and interact with structural cells in airways and lung parenchyma.

The wide variety of inflammatory mediators that have been shown to be increased in COPD patients [36] attract inflammatory cells from circulation amplifying the inflammatory process and inducing structural changes that may lead to emphysema and mucus hypersecretion [37].

Lung inflammation is believed to be further augmented by oxidative stress and an excess of proteinases in the lung. These two mechanisms are key players in COPD pathology (Figure 1).

It has long been proposed that several proteases disrupt connective tissue components, as elastin, in lung parenchyma to produce emphysema and that there is an imbalance in COPD patients between proteases and endogenous antiproteases which should protect the lung against protease-derived effects. In COPD, the exogenously and endogenously derived oxidants have been found to inactivate antiproteinases such as α1-antitrypsin [38]. Evidences of elastin degradation in COPD have been demonstrated and although early attention was directed to neutrophil elastase, many other proteases have been reported to be able to degrade elastin [7].

Oxidants are generated endogenously and exogenously, with cigarette smoke being heavily implicated in the latter as it contains many oxygen free radicals [31]. Under normal circumstances and despite permanent exposure to high oxygen levels, the lung is able to manage oxidant species by neutralizing them with several antioxidant mechanisms in the human respiratory tract [7,31,39]. Oxidative stress occurs when reactive oxygen species (ROS) are produced in excess of the antioxidant defense mechanisms resulting in harmful effects such as damage to lipids, proteins and deoxyribonucleic acid (DNA) [7,40]. Inflammatory and structural cells that are activated in the airways of COPD patients produce ROS, including neutrophils, eosinophils, macrophages and epithelial cells [7,40,41]. Alveolar macrophages are activated by free radicals and react by producing high levels of mediators, some of which are chemotactic for neutrophils and macrophages, as well as ROS and also reactive nitrogen species (RNS), with resultant local and systemic inflammation [31].

It is increasingly recognized that the inflammatory response associated with COPD extend beyond the lung [42]. Evidence of systemic inflammation includes activated circulating inflammatory cells and elevated levels of both inflammatory cytokines and acute phase proteins as C-reactive protein, fibrinogen, leukocytes and tumor necrosis factor (TNF-α) in COPD patients when compared to healthy subjects [43].

The origin of systemic inflammation in COPD is still unclear and requires further investigation, but it is likely to be a consequence of a number of factors, including individual susceptibility and the direct effects of hypoxia and noxious substances as the one of cigarette smoke on the peripheral vasculature and circulating inflammatory cells [31]. Alternatively, the observed inflammation may be a consequence of ‘overspill’ from the lung to the peripheral circulation [31]. Systemic inflammation is directly linked to a number of complications commonly encountered in COPD patients including, but not limited to, cachexia, skeletal muscle dysfunction, depression, osteoporosis, diabetes/glucose intolerance, autoimmune disorders and cardiovascular diseases [31,42].

Among respiratory diseases, COPD is the leading cause of lost work days. In the United States of America, medical costs credited to COPD were estimated at $32.1 billion [44]. In the European Union, productivity losses are estimated to amount to a total of €28.5 billion annually [12]. Economic burden is likely to be underestimated since, for example, the economic value of the care provided by family members is not generally acknowledged. Long-term home care provided by relatives for COPD patients has a negative impact on professional careers for both patients and their family members [5]. Hence, COPD represents a very important threat to global economies and the development of new diagnostic/prognostic methodologies capable of providing exact medical information to the doctor in a timely manner is crucial to diminish the economic impact of COPD worldwide.

Proteomics
Definition
The Proteome is, by definition, the total set of proteins expressed by a given cell, tissue, organ or organism at a certain time under certain defined conditions [45]. Proteomics is defined by the large-scale study of the proteome. The human genome has been sequenced a decade ago and about 20,000 genes were accounted. Genome sequencing contains valuable information to proteomics. Proteomics has the ability to raise our knowledge to a higher level providing new insights into the pathophysiology of many diseases ultimately leading to new prognosis and diagnosis, and also to novel therapeutical treatments.

Background and state-of-the-art
The advent of proteomics has brought with it the hope of discovering novel biomarkers that can be used to diagnose a specific disease, predict susceptibility and monitor disease progression and treatment responses, among many other applications. This hope is built on the ability of proteomic technologies, such as mass spectrometry (MS), to identify hundreds of proteins in complex biofluids such as plasma and serum. Very few if any analytical instruments surpass the mass spectrometer in the versatility of its application in both basic and applied research, as it is the case of biomarker discovery. To support this statement, it is sufficient to mention that mass spectrometry can be used for applications ranging from characterization of electronic excited states and vibrational levels of simple molecules to the construction of protein interaction maps in multicellular organisms. This is also the result of almost one hundred years of mass spectrometry utilization since Sir J. J. Thomson was able to create the first mass spectrometer in 1913 [46]. For over 80 years ionization methods have excluded the study of large molecules, including peptides and proteins. In the 1980s, this paradigm changed with the introduction of new ionization methods as electrospray ionization (ESI) [47] and matrix-assisted laser desorption ionization (MALDI) [48,49]. These simple and sensitive ionization methods have been coupled to different types of analyzers such as triple quadrupoles, three-dimensional ion traps, and time of flight (TOF), including its orthogonal version which allowed coupling of TOF to both pulsed (MALDI) and continuous (ESI) ionization types [45]. A further impetus was given to the process of ion analysis through the commercialization of hybrid configurations that have been intensively used in proteomics including, but not limited to, TOF-TOF, ion trap-Fourier transform (FT)-ion cyclotron resonance (ICR), and quadrupole-TOF. These combinations have a direct impact on sensitivity and resolution of the sequence information that can be obtained when performing tandem MS analysis [45]. At the same time MS was evolving, there were many advances in other fields that were crucial to the development of proteomics as sample preparation techniques and bioinformatics tools. One of the most widely used separation procedures is two-dimensional gel electrophoresis (2DE) which consists in the separation of a complex protein mixture
according to physicochemical properties of proteins. First, proteins are separated in one dimension according to their isoelectric point through isoelectric focusing (IEF) in immobilized pH gradient (IPG) strips, and then separated over a second dimension according to their molecular weight in a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). This separation method was described as it is used today in 1975 by O’Farrell [50]. SDS-PAGE is one of many achievements that took place in the last 50 years and that established proteomics in the first line of clinical research at the present time [51].

In last decade, the shotgun proteomics approach has become the method of choice for identifying and quantifying proteins in most large-scale studies [52-55]. Compared with 2DE, shotgun proteomics allows higher data throughput and better protein detection sensitivity. This strategy is based on digesting proteins (usually with trypsin) into peptides. This produces a complex peptide mixture that is then separated by one- or multiple dimensional liquid chromatographies (LC) and subjected to peptide sequencing using tandem mass spectrometry (MS/MS) before automated database searching. This strategy is compatible with the use of labeled samples for quantitative purposes such as stable isotope labeling by amino acids in cell culture (SILAC) [56], isotope coded affinity tags (ICAT) [57], isobaric tags for relative and absolute quantitation (iTRAQ®) [58] or by O¹⁵/O¹⁷ exchange [59].

The chance of combining different techniques along sample preparation steps with different separation methods and different types of mass spectrometers generates multiple complementary approaches whose results can be combined to achieve a higher level of understanding (Figure 2). Proteins are the functional molecules in the human body and therefore by accurately assessing the proteomes of healthy and disease states (including information on protein differential expression across conditions) we will be one step closer to determine what changes are responsible for disease onset and progression, to properly determine disease stages and to be able to monitor treatment efficacy.

**Figure 2:** Workflow illustrating different proteomics-based approaches and major steps required for proteomic biomarker discovery.

**Proteomics in COPD**

There is still some debate concerning the disease-specific molecular mechanisms of the inflammatory process and acute exacerbation of COPD. Potential biomarkers which are specific for COPD have not been fully identified and validated, even though there is a great need for such biomarkers [60]. Proteomic technologies allow for identification of protein changes caused by the disease process and recent advances, especially at mass spectrometry and bioinformatics levels, raise the chances to identify novel putative biomarkers. Proteomics application to clinical samples is extremely challenging as it is subjected to natural variation of the human proteome (geneic, environmental) and also to technical aspects prior to analysis as sample collection, storage, handling and processing procedures. In particular, COPD proteomics faces the huge challenge of sample collection since obtaining enough amounts of proper samples is dependent on factors such as ethics and accessibility. Biological samples commonly used in COPD include bronchoalveolar lavage fluid (BALF) and sputum, but exhaled breath (EB), nasal lavage fluid (NLF), urine, plasma and bronchial and lung biopsies have also been used to study the disease.

Despite its potential to investigate the mechanisms of many diseases and the existence of a systemic component in COPD, the use of plasma had limited importance to COPD proteomics so far. A similar rationale can be stated for lung and bronchial biopsies. These samples possess enormous potential for determining the agents and mechanisms at the site disease and therefore to identify biomarkers that could then be translated into the clinic. However, the procedure to collect these samples is extremely invasive and collected samples often differ in terms of histology and protein content. Sputum is considered to represent bronchial lining fluid and induced sputum may in this way be an important source for proteomics research since it contains enough protein amounts to engage proteomics studies. BALF and NLF result from respiratory epithelial lining washing, which are relatively non-invasive techniques that are able to collect samples which are representative of the upper airways. Our group has successfully demonstrated that the use of nasal epithelial cells (NEC) collected by
nasal brushing is capable of yielding numerous and well-preserved dissociated cells that are representative of the human superficial respiratory mucosa [61] and their utility in the study of the monogenic disease cystic fibrosis by proteomics [62,63]. NECs were also reported to constitute an accessible surrogate for studying lower airway inflammation [64]. Importance of these cells and their application to respiratory proteomics have recently been reviewed [65].

Information on putative biomarkers for COPD generated from several proteomics studies has been gathered in recent reviews [60,66]. Samples collected form BALF, sputum and lung tissues were able to provide differentially expressed proteins in COPD. Regarding BALF, there were four proteins reported to be differentially expressed in COPD (Neutrophil defensins 1 and 2 and calgranulin A and B) [67] and in sputum, Clara cell secretory protein (CCSP) and SP-A were the two proteins reported as potential biomarkers [68,69]. Concerning lung tissues, two studies mentioned in this review yielded 12 differentially expressed proteins when compared to healthy controls, including matrix metalloproteinase-13 (MMP-13) and again surfactant protein A (SP-A) [69,70].

Surprisingly, to date only 57 reports (14 are reviews) match the search at Pubmed (http://www.ncbi.nlm.nih.gov/pubmed, accessed December 20, 2011) for COPD proteomics, while proteomics and COPD account for about 30,000 each when separately searched. Hence, there is a clear need to engage clinically valuable proteomics studies in order to match the need for new biomarkers in COPD. Proteomics has proved to be an excellent tool to identify disease biomarkers in COPD and sample collection is indeed what needs to improve. Accurate and updated clinical data ranging all aspects directly and indirectly related to the disease is the most important step in clinical proteomics and it is particular important to COPD since COPD is a combination of distinct features, including the systemic ones. Then, the source of biospecimens is also of extreme importance, not its choice, but especially to what standard collecting procedures is concerned. This may involve training clinicians in biospies collection to ensure reproducibility and adoption of a standard procedure for sample handling and processing until the results are generated. To overcome ethics and invasiveness issues, NECs collected by the non-invasive and painful procedure of nasal brushing may be the right option to study COPD by means of proteomics. Another important issue is that COPD shares some clinical features with other pulmonary diseases such as asthma or cystic fibrosis and therefore comparing the proteome of COPD patients with healthy controls may not be enough when searching for specific biomarkers for the disease. It is recommended to incorporate patients suffering from other disorders that mimic features of COPD as control groups for discovery of reliable specific biomarkers for COPD [71].

Once the aforementioned premises are fulfilled, proteomics and COPD will certainly be a match for success and coming years will confirm this statement and clinicians and their patients will greatly profit from this alliance.

References
33. Lapperre TS, Postma DS, Gosman MM, Snoeck-Stroband JB, ten Hacken NH,