Biomarkers, 2010; 15(4): 367–377

informa
healthcare

ORIGINAL ARTICLE

High-resolution mass spectrometry proteomics for the identification of candidate plasma protein biomarkers for chronic obstructive pulmonary disease

Timothy P. York[1,2], Edwin J. C. G. van den Oord [2], Timothy B. Langston[3], Jeffery S. Edmiston[3], Willie McKinney[3], Bradley Todd Webb[2,4], E. Lenn Murrelle[4], Barbara K. Zedler[4], and Jason W. Flora[3]

[1]Department of Human and Molecular Genetics, [2]Center for Biomarker Discovery and Personalized Medicine, Virginia Commonwealth University, Richmond, VA, USA, [3]Altria Client Services, RD&E, Richmond, VA, USA, and [4]Venebio Group, LLC, Virginia Bio-Technology Research Park, Richmond, VA, USA

Abstract

Although cigarette smoking is recognized as the most important cause of chronic obstructive pulmonary disease (COPD), the pathophysiological mechanisms underlying the lung function decline are not well understood. Using off-line strong cation exchange fractionation with RP-LC-ESI-MS/MS and robust database searching, 1758 tryptic peptides were identified in plasma samples from cigarette smokers. Using two statistical approaches, 30 peptides were identified to be associated with the annualized rate of lung function decline over 5 years among smokers with COPD characterized as having rapid ($n = 18$) or slow ($n = 18$) decline and 18 smokers without COPD. The identified peptides belong to proteins that are involved in the complement or coagulation systems or have antiprotease or metabolic functions. This research demonstrates the utility of proteomic profiling to improve the understanding of molecular mechanisms involved in cigarette smoking-related COPD by identifying plasma proteins that correlate with decline in lung function.

Keywords: *COPD; cigarette smoking; biomarker; mass spectrometry; protein; plasma*

Introduction

Chronic obstructive pulmonary disease (COPD) is projected to be the third leading cause of death in the world by 2020 (Ezzati & Lopez 2003, Mannino et al. 2002, Pinto-Plata et al. 2007), and cigarette smoking is widely recognized as its primary causative factor. The pulmonary component of COPD is primarily characterized by chronic airway inflammation and incompletely reversible, usually progressive, airflow obstruction (Barnes et al. 2003, Rabe et al. 2007). The operational diagnosis of COPD has traditionally been made by spirometry, as a ratio of the forced expiratory volume in one second (FEV_1) to the forced vital capacity (FVC) below 70% (Rabe et al. 2007). Pathophysiological mechanisms believed to underlie COPD include an imbalance between proteinase and antiproteinase activity in the lung, dysregulation of antioxidant activity and chronic abnormal inflammatory response to long-term exposure to noxious gases or particles leading to the destruction of the lung alveoli and connective tissue (Barnes et al. 2003, Rabe et al. 2007). However, COPD is increasingly recognized as a syndrome associated with significant systemic effects which are attributed to low-grade, chronic systemic inflammation (Agusti et al. 2003, Agusti & Soriano 2008, Fabbri & Rabe 2007, Rahman et al. 1996).

The aim of this study was to improve the understanding of the molecular mechanisms involved in cigarette

Address for Correspondence: Timothy P. York, Department of Human and Molecular Genetics, Virginia Commonwealth University School of Medicine, PO Box 980003, Richmond, Virginia, 23298-0003, USA. Tel: +1 804 828 8757. Fax: +1 804 827 0018. E-mail: tpyork@vcu.edu

(Received 09 January 2010; revised 18 March 2010; accepted 18 March 2010)

ISSN 1354-750X print/ISSN 1366-5804 online © 2010 Informa UK Ltd
DOI: 10.3109/13547501003789901

smoking-related COPD by identifying plasma proteins that correlate with decline in lung function. With the recent advancements in protein and peptide separation, coupled to highly selective (e.g. high resolution) and sensitive mass spectrometry, proteomic techniques are an excellent platform for the discovery of biomarkers of disease and disease progression. Proteomic profiling to facilitate biomarker discovery and disease mechanistic insight has shown considerable potential in lung disease (Jacot et al. 2008) and in other areas of research, for instance, prenatal diagnostics (Choolani et al. 2009), prostate cancer (Rosenzweig et al. 2009) and heart disease (Fu & Van Eyk 2006, Vivanco et al. 2005). An ideal biomarker is found in biosamples such as plasma, urine or saliva that can be readily collected using minimally or non-invasive methods (Meng & Veenstra 2007). Of these biofluids, plasma is probably the most well known; however, it is also the most complex. Several proteomics approaches have been employed to elucidate the plasma proteome, including separation techniques such as two-dimensional gel electrophoresis, intact protein fractionation, peptide fractionation, and nanoflow liquid chromatography (LC). Mass spectrometry (MS) is almost exclusively the detection technique using either matrix-assisted laser desorption ionization (MALDI) or electrospray ionization (ESI). The MS platforms by which these measurements are made typically involve time-of-flight (TOF), ion-trap, and Fourier transform ion cyclotron resonance MS (FTICR) or hybrid instruments such as quadrupole-TOF and an ion-trap-FTMS. Recent observations by the Human Proteome Organization's (HUPO) Plasma Proteome Project have shown that off-line peptide separation by strong cation exchange (SCX) followed by reversed-phase (RP) LC with ESI-MS/MS will result in the identification of more proteins of low abundance (Li et al. 2005).

In this study we compared differential protein expression among two groups of adult cigarette smokers with mild to moderate COPD but different rates of lung function decline and a gender-matched group of smokers without COPD. Identification of proteins that are differentially abundant among the groups may allow a better understanding of the mechanisms underlying cigarette smoking-related lung function decline.

Materials and methods

Subjects

Subjects were selected from the 244 University of Utah study centre participants in the Lung Health Study (LHS) who also participated in the follow-on – the Genetics of Addiction Project (GAP). LHS enrolled male and female cigarette smokers, aged 35–60 years, with mild or moderate COPD, in a prospective, randomized, multicentre clinical study (Anthonisen et al. 1994). GAP was a cross-sectional assessment which also enrolled 94 adult cigarette smokers without COPD as a control group. Smoking status was assessed and lung function was measured by spirometry at baseline (1986–89), annually for 5 years, once during 1998–2001 (Anthonisen et al. 2002), and once in GAP (2003–04). Spirometry included FEV_1 and FEV_1 adjusted for age, sex and height (i.e. as percentage of predicted) (Rabe et al. 2007). The annualized rate of lung function decline during the 5 years of LHS was calculated for each participant as the slope of the linear regression of FEV_1%.

The costs associated with MS/MS protein quantification constrained the number of samples we could process; therefore we selected phenotypic extremes to maximize the biological variability in our study (Navabi et al. 2009). A subset of 54 GAP participants was selected for plasma proteomic analysis in this study: the 18 with the steepest rate of decline in FEV_1 (rapid decliners, RPD), the 18 with the least steep or no annualized rate of decline in FEV_1 (slow decliners, SLW), and 18 smokers without COPD as a control group. Characteristics of the three groups are shown in Table 1. Over the first 5 years of LHS, the rapid decliners had an average annual decrease in FEV_1 of 1.6% predicted/year while the slow decliners had an average increase of 0.8% predicted/year. At the GAP assessment approximately 17 years after baseline, 7/18 (39%) of the RPD participants and 12/18 (67%) of the SLW participants had quit smoking, and in the control group, 8/18 (44%) had quit smoking in the 3 months before GAP participation ($\chi^2 = 3.11$, 2 d.f., $p = 0.21$).

Plasma sampling and processing

Plasma was sampled at the GAP assessment by venipuncture using a sodium citrated Vacutainer® tube at least 2 h after eating. Within 10 min of collection, blood was centrifuged for 15 min at 1500*g* and 2–6°C. The topmost plasma was removed and further centrifuged at 1500*g* for 15 min. Plasma samples were shipped on dry ice, stored at -80°C, and thawed just before analysis.

Sample pooling

In each of the three study groups, plasma samples from six subjects were pooled to reduce heterogeneity within the group, increase yield of group-specific low-abundance proteins/peptides and minimize instrument run time. Therefore, three pools were evaluated for each of the three study groups, for a total of nine plasma sample pools. Samples were selected for each pool by applying a random number generator.

Depletion of high-abundance plasma proteins

All pooled plasma samples were depleted of the top 12 most abundant proteins using a Beckman Coulter

Table 1. Characteristics of study participants.

Characteristic	Cigarette smokers with COPD,rapid decline[a] (RPD) ($n=18$)			Cigarette smokers with COPD,slow decline[a] (SLW) ($n=18$)			Cigarette smokers without COPD ($n=18$)	
	Lung Health Study			Lung Health Study				
	Baseline	Year 5	GAP	Baseline	Year 5	GAP	GAP	*p*-Value
Male, *n* (%)			13 (72.2)			10 (55.6)	9 (50.0)	0.369[b]
Age (years), mean (SD)			64.8 (5.4)			63.6 (7.3)	57.2 (7.7)	0.002[c]
Cigarettes per day, mean (SD)[e]	37.3 (17.2)	21.6 (19.0)	16.0 (17.0)	27.7 (10.7)	8.3 (11.4)	5.6 (9.5)	9.7 (12.0)	0.167[c]
Years smoked, mean (SD)			42.1 (6.8)			34.8 (9.2)	30.1 (11.3)	<0.001[c]
FEV_1 (L), mean (SD)	2.75 (0.59)	2.34 (0.67)	1.70 (0.60)	2.61 (0.57)	2.63 (0.66)	2.32 (0.55)	3.20 (0.63)	
Δ FEV_1 (L), mean (SD)		-0.40 (0.23)			0.02 (0.21)		na	<0.001[d]
FEV_1 % predicted, mean (SD)	76.1 (9.7)	67.8 (13.3)	54.6 (16.6)	74.8 (9.7)	78.5 (12.1)	77.2 (14.2)	103.1 (18.4)	
Δ FEV_1 % predicted, mean (SD)		-8.22 (7.34)			3.77 (5.73)		na	<0.001[d]

COPD. chronic obstructive pulmonary disease; GAP, Genetics of Addiction Project, an average of 17 years after baseline at which time plasma proteomic analysis was performed; FEV_1, forced expiratory volume in 1 s; na, not applicable.

[a]Decline in lung function was assessed as the slope of a linear regression of the annualized rate of decline during the first 5 years of participation in the Lung Health Study in FEV_1 % predicted (i.e. adjusted for age, height, and gender). [b]$\chi^2 = 1.99$, 2 d.f. test. [c]Test of association between characteristic and lung function at GAP by linear regression. [d]Change in characteristic from baseline to year 5 for RPD versus SLW. [e]At the GAP time point, 7/18 (39%) of RPD, 12/18 (67%) of SLW, and 8/18 (44%) of control subjects had quit smoking; $\chi^2 = 3.11$, 2 d.f., $p = 0.21$.

IgY-12 High Capacity spin column (part #A24618) using the recommended manufacturer's procedure. In short, 20 µl of plasma were added to 480 µl of dilution buffer. The samples were then filtered through 0.22 µm spin filters by centrifugation for 1 min at 16 000*g*. The depletion columns were then centrifuged for 30 s at 400*g* to dry the beads. The end caps were attached and the diluted plasma sample was added and mixed by inverting the column. The sample was placed on a rotator (end to end) and incubated at room temperature for 30 min. Columns were then inverted and the tips were removed. The samples were then placed in collection tubes and centrifuged for 30 s at 400*g*. The flow-through was then collected for concentration and digestion. The bound proteins (high abundant top 12) were then stripped from the beads with three consecutive washes and centrifugations. The flow-through was discarded and the spin columns were regenerated (manufacturer's procedure).

Protein digestion (plasma)

The depleted flow-through was added to a pre-rinsed Microcon YM-3 (3000 Da) molecular weight cut-off spin cartridge (Millipore), following the manufacturer's recommended protocol, and centrifuged at 14 000*g* until 100 µl of retentate remained (~30 min.). The retentate was then transferred to a clean microcentrifuge tube and proteins were reduced using 15 µl of 50 mM ammonium bicarbonate (Pierce) and 1.5 µl of 100 mM DL-1,4-dithiothreitol (Acros, Geel, Belgium) and incubation at 95°C for 5 min. After samples had cooled, they were alkylated by the addition of 3 µl of 100 mM iodoacetamide (Pierce) and were incubated for 20 min in the dark at room temperature; 1.5 µl of 100 ng µl^{-1} porcine trypsin (Promega, Madison, WI, USA) was then added and the samples were incubated at 37°C for 3 h. An additional 1.5 µl of 100 ng µl^{-1} trypsin was then added followed by incubation at 37°C for approximately 16 h. To ensure sufficient reagent mixing, all samples were vortexed (30 s) and centrifuged (2000*g* for 1 min) following each solution addition. Samples were dried in a vacuum centrifuge at 45°C. Samples were reconstituted with 50 µl of 3% acetonitrile with 0.1% formic acid and vortexed (30 s) prior to fractionation.

Off-line plasma fractionation

Off-line fractionation of the plasma tryptic digests in each pooled sample was conducted using a GE Healthcare MDLC Ettan (Piscataway, NJ, USA) fitted with a GE FRAC950 fraction collector fitted with a strong cation exchange (SCX) column (Thermo Fisher Scientific Biobasic SCX, 250 × 2.1 mm). Ion exchange (IEX) solvent A was 20 mmol l^{-1} citric acid (Fisher) in 75% HPLC grade water and 25% acetonitrile (Fisher) (3.8 g citric acid in 1 l of 25% acetonitrile) (pH 2.65). IEX solvent B was 20 mmol l^{-1} citric acid and 1 mol l^{-1} ammonium chloride (Fisher) in 75% HPLC grade water and 25% acetonitrile (3.8 g citric acid and 53 g ammonium chloride dissolved in 1 l 25% acetonitrile, pH 2.65). The fraction collector was conditioned for approximately 20 min before each run with 100% IEX solvent A at 200 µl min^{-1}. The tryptic digest plasma samples were reconstituted in 50 µl of IEX solvent A. Run parameters began a 40 µl sample injection and 0% IEX solvent B for 10 min, ramped to 60% IEX solvent B in 30 min, then to 100% IEX solvent B and held for 5 min. The system flow rate was 200 µl min^{-1} and fractions were collected each minute in a 96-well plate (200 µl fractions). For each sample, the ten 200 µl fractions that contained peptides (determined in method

development) were lyophilized at 45°C and stored at -20°C until analysis. Samples were reconstituted with 50 µl of 3% acetonitrile with 0.1% formic acid and vortexed (30 s) prior to analysis.

Liquid chromatography

All ncap-LC was conducted using an Eksigent nanoLC-1D (Monmouth Junction, NJ, USA) with a Leap Technologies (Carrbaro, NC, USA) autosampler and a Zorbax 300SB-C8 trap column (5 × 0.3 mm). Reverse-phase separation was conducted on each of the ten fractions from each pooled sample using a New Objective Picofrit Proteopep™2 (5 cm of C18 packing and a 15 µm tip). The LC run program had a 4-min trap wash at 10 µl min^{-1}, a 10 µl injection volume and a 270 nl min^{-1} flow rate. LC buffer A contained 0.1% formic acid in LCMS grade water (Fisher) and B contained 84% high purity acetonitrile (Fisher) with 0.1% formic acid. The LC gradient started at 3.5% B and ramped to 9% B in 1 min. The gradient was ramped to 70% B in 37 min, 97% B for 12 min and then returning to 3.5% B.

Mass spectrometry

All data were collected on a Thermo-Finnigan (San Jose, CA, USA) LTQ-FTMS (a hybrid linear ion-trap with a 7 Tesla FTICR MS) with Xcalibur™ 2.0 and fitted with a New Objective Picoview 550 nanospray ionization source. Full scan data were collected at 50 000 resolution (at 400 *m/z*) in the FTMS with a mass-to-charge ratio (*m/z*) range of 400–2000. The instrument was externally calibrated no less than 5 days prior to acquisition following manufacturer's recommended protocol with caffeine, MRFA and Ultramark. All data were collected using data-dependent scanning with multistage MS (MS/MS) using collision-induced dissociation (CID) with a 3 *m/z* isolation width, normalized collision energy of 35 and 30 ms activation in the ion-trap MS (LTQ, unit mass resolution) on the top five most abundant peaks. Charge state screening and monoisotopic precursor selection were enabled. The acquisition had a 30 s dynamic exclusion using an *m/z* range of 0.01 low to 1.01 high for the exclusion list with an exclusion limit of 500 *m/z* values.

Database searching

Database searching was conducted using Thermo-Finnigan Bioworks 3.3.1 SP1. The Human Refseq database was used (download November 2007) for all searches. The number of entries in the Refseq database was 34 180. Prior to the SEQUEST search, the Human Refseq database was indexed for trypsin (KR), monoisotopic mass, fully enzymatic (cleavage at both sides), molecular weight range of 400–10 000, three missed cleavage sites, variable modifications with post-translational modifications of oxidation of the methionines at 15.99492 Da and alkylation of the cysteines at 57.02146 Da. Mass accuracy was set to 20 parts per million for precursors and 1 AMU for the fragment ion tolerance. For all fractions of each pool, individual SEQUEST files were combined using the Bioworks Multiconsensus report function. The rigorous SEQUEST search constraints were set with a Delta CN ≥0.100 and Xcorr vs charge state of 1.9 for 1+, 2.2 for 2+, and 3.75 for 3+ as suggested by the Human Proteome Organization (HUPO) (Omenn et al. 2005) and 4.0 for 4+. The number of different peptides allowed for protein identification was set to one. The total peak areas were determined using the Bioworks algorithm PepQuan with parameters set to area, mass tolerance of 0.0100, minimum threshold of 1000, number of smoothing points at 5, and including all proteins. The false discovery rate was estimated to be less than 10%. Briefly, a concatenated target-decoy database was created using the human Refseq database. Results were searched against the concatenated database and false positives were estimated as twice the number of passing decoy fragments. The false discovery rate was determined by dividing the false positives by the sum of the true positives and false positives (Elias & Gygi 2007).

Statistical analysis

Despite extensive quality control efforts to maintain a high level of data reproducibility, results from independent peptide analyses of plasma pools will vary in the distribution of peptide abundance values due to expected variability in the experimental process. Thus to allow for comparisons across pools, the median-centred natural logarithm of peptide abundance (peak area) within each pool was calculated to standardize abundance values. We undertook two approaches to handle the large amount of missing data which is typical for MS/MS-based proteomic studies. In the first case, assuming missing data represent abundance values below the detection threshold, data were imputed to a value of one-half the minimum intensity for each pool plus a small amount of random error. In the second case, missing data were not imputed and thus no assumptions were made about the source of missing data, such as technical error or the real absence or low abundance of protein in plasma. To identify peptides correlated linearly with the presence of COPD and an increasing rate of lung function decline, the study groups were coded ordinally (control = 1, SLW = 2 and RPD = 3) and regressed against the standardized peptide abundance values. Peptides were included (filtering constraints) if observed in at least three of the nine sample pools. In the non-imputation method, we further required peptide presence in each of the three study groups (filtering constraints). As the condition of normality of each peptide predictor in the linear regression

model cannot be guaranteed, we obtained empirically derived *p*-values by a permutation test with 1000 iterations. We corrected for multiple testing by calculation of the false discovery rate (FDR) and reported the corresponding *q*-values (Benjamini & Hochberg 1995, Storey & Tibshirani 2003a, b, van den Oord 2005).

Protein annotation and pathway analysis

Mapping of proteins to pathway was conducted on Kyoto Encylopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/) (Kanehisa et al. 2006, 2008, Kanehisa & Goto 2000). Pathway analysis was conducted with Pathway Studio™ software version 5.0 from Ariadne Genomics, Inc. (Rockville, MD, USA). The analysis was manually filtered using the expanded pathway analysis tool and limiting analysis to proteins.

Results

Figure 1 shows an example of the data analysed in this study and demonstrates the high complexity and quality of the data acquired by the hybrid ion-trap-FTMS. This figure shows the total ion chromatogram for fraction 7 of one of the three sample pools in the RPD group. It includes the MS spectrum at 35.92 min and zooms in on the peptide with the amino acid sequence

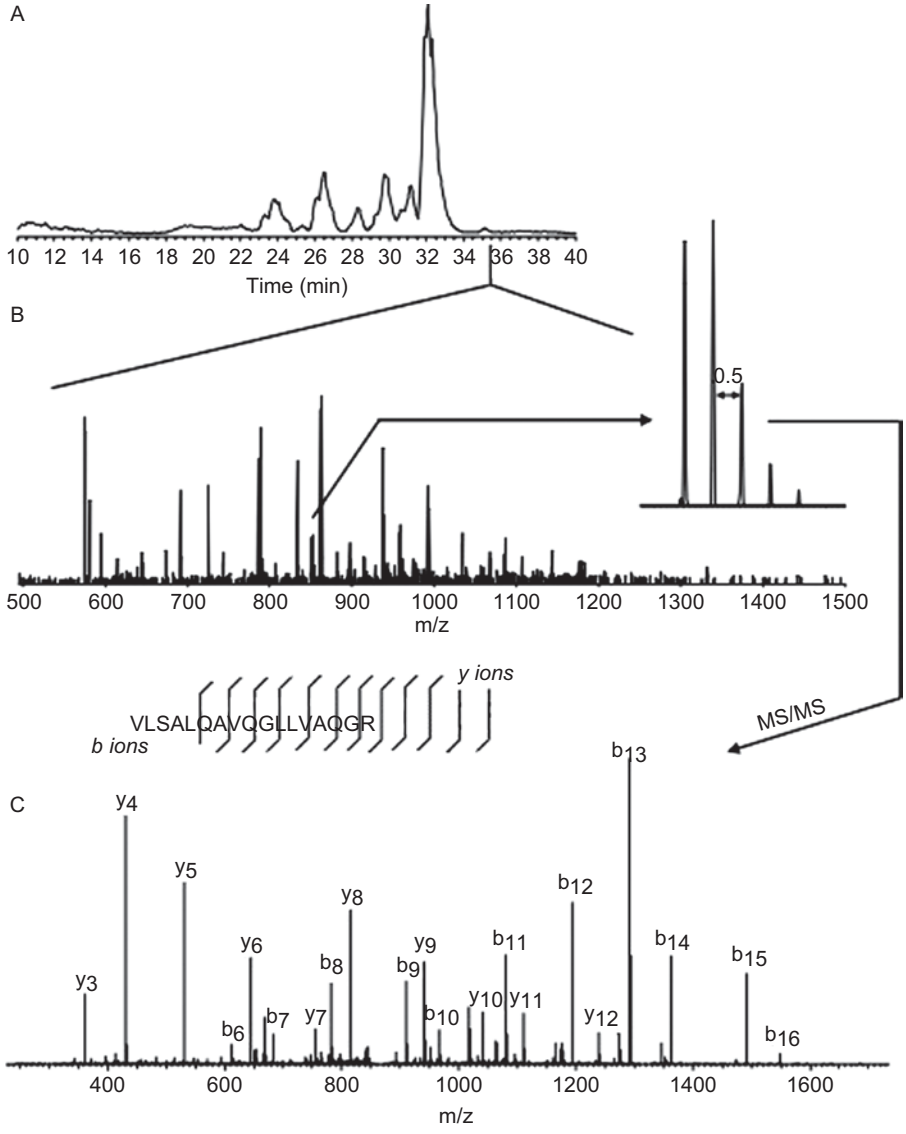

Figure 1. (A) Total ion chromatogram for fraction 7 of a rapid decline chronic obstructive pulmonary disease (COPD) plasma pool. (B) Full-scan FTMS spectrum at 35.92 min plus the zoom in on FTMS spectrum showing the isotopic fine structure of the tryptic peptide of human angiotensin with the amino acid sequence VLSALQAVQGLLVAQGR found at *m/z* 862.01514 with a 2+ charge state and <5 ppm mass accuracy. (C) Fragmentation spectrum for *m/z* 862 with y and b ions labelled and the peptide bond cleavages noted in the peptide sequence.

VLSALQAVQGLLVAQGR found at *m/z* 862.01514, which maps to human angiotensin. This peptide was detected at the 2+ charge state with a mass accuracy <5.0 ppm using external calibration. The fragmentation data (MS/MS) is shown and 'y' and 'b' ions are labelled showing significant coverage of the amino acid sequence.

Off-line SCX fractionation with RP-LC-ESI-MS/MS and robust database searching resulted in the observation of 1758 unique peptides across all nine pooled samples. The filtering constraints (discussed in the methods section) for the imputation and non-imputation methods resulted in 1133 and 973 peptides, respectively, for statistical analysis. At an FDR level of 10%, a total of 17 peptides were significantly associated with lung function decline for the imputation method, 20 peptides were significant for the non-imputation method and seven of these peptides were identified by both methods (Table 2). The regression coefficients from the linear model, along with the associated *q*-value for each method where applicable, are also presented in Table 2 for each unique peptide. A negative regression coefficient estimate indicates linearly decreasing peptide abundance levels across the three study groups, from controls to SLW to RPD, while a positive estimate indicates a linear increase in peptide abundance levels from controls to RPD.

The 30 unique peptides identified as differentially expressed across the three study groups by linear regression mapped to 21 unique proteins. In Table 2 the peptides are grouped according to major function. The majority of the identified peptides (17/30), representing 12 proteins, are involved in the complement cascade which, as part of the innate immune system, promotes host defence mechanisms of bacterial lysis, phagocytosis and immune cell recruitment and activation (Bolger et al. 2007, Markiewski & Lambris 2007). Regression analysis across the three study groups indicated a mixed pattern of over- and underexpression among the 17 complement-related peptides.

A peptide unique to SERPINA3, or serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3, was relatively underexpressed in the RPD group compared with the SLW group, and highest levels were in the control group. SERPINA3 is a protease inhibitor and lower levels of this protein in the RPD support an imbalance in proteases/antiproteases in the RPD population. Certain allelic variants of *SERPINA3* result in reduced protease inhibitor activity and have been associated with COPD (Hersh et al. 2008, Ishii et al. 2000, Poller et al. 1992).

Peptides from fibrinogen, kallikrein B and inter-alpha (globulin) inhibitor H1, all components of the coagulation system, were relatively overexpressed in the COPD groups compared with the control group, with the highest levels in the RPD group. Coagulation is a complex cascade involving plasma proteins and platelets that results in blood clot formation (Furie & Furie 2005). Circulating clotting factors and their proteases and antiproteases regulate this process (Fay et al. 2007) and the coagulation system is thought to be involved in the thromboembolic complications associated with COPD and smoking (Tapson 2005, Voelkel & Cool 2003, Yanbaeva et al. 2007). Peptide abundances also suggest that plasma levels of antithrombin (SERPINC1), a component of one of the principal intrinsic anticoagulant systems (Rosenberg & Bauer 1987), and alpha-2 antiplasmin, a major regulator of intravascular fibrinolysis (Carpenter & Mathew 2008), were highest in the control group and lowest in the rapid FEV_1 decline COPD group.

Alpha-2 antiplasmin is also involved in the renin–angiotensin system (RAS) as a critical regulator of angiotensin II-mediated vascular remodelling (Huo 2008). Angiotensinogen is an inactive circulating substrate which is converted by renin to angiotensin I, the precursor peptide in the classical RAS cascade (Ribeiro-Oliveira et al. 2008). The circulating and local tissue RAS are involved in vascular remodelling (Kumar et al. 2008) and play pivotal pathophysiological roles in hypertension (Lalouel et al. 2001) and diabetes (Ribeiro-Oliveira et al. 2008). Experimental evidence suggests that oxidant stress-induced damage of lung microvascular endothelial cells in cigarette smokers results in endothelial cell apoptosis, capillary loss, impaired angiogenesis and profound airspace enlargement (Voelkel et al. 2007).

Peptide abundances indicate that the expression of two insulin-like growth factor binding proteins (IGFBP3 and IGFALS) was lowest in the RPD COPD group and highest in the control group. This possibly reflects the lower levels of anabolic hormones, such as insulin-like growth factors and testosterone, found in chronic inflammatory muscle-wasting conditions such as COPD, chronic heart failure, acquired immunodeficiency syndrome and cancer (Sevenoaks & Stockley 2006, Wouters et al. 2002).

Three peptides mapping to apolipoprotein B100 had a mixed pattern of differential expression across the study groups. Apolipoprotein B is the major structural protein of very low- and low-density lipoproteins (VLDL, LDL), and apoB-containing lipoproteins transport cholesterol from the liver and gut to peripheral tissues (Marcovina & Packard 2006). On the other hand, apolipoprotein A-IV is the major protein component of high-density lipoproteins (HDL) which reverse transport cholesterol from the periphery to the liver for excretion (Marcovina & Packard 2006) and a potent endogenous inhibitor of lipid oxidation (Qin et al. 1998). A peptide from apolipoprotein A-IV suggests that this protein was relatively underexpressed in the COPD RPD group compared with the COPD SLW and control groups.

Gelsolin (GSN) is an actin-binding protein involved in regulating host response to cellular damage in bacterial sepsis (Lee et al. 2007). Two peptides with opposite directions of differential expression across the study groups

Table 2. Unique peptides differentially expressed across the three study groups: smokers without chronic obstructive pulmonary disease (COPD), smokers with COPD with slow forced expiratory volume in 1 s (FEV_1) decline (SLW), and smokers with COPD with rapid FEV_1 decline (RPD). Reported regression coefficient estimates for each peptide were significant at the 10% false discovery rate.

			No imputation		Imputation	
Peptide sequence	Gene symbol	Name	Regression coefficient[a]	*q*-Value	Regression coefficient[a]	*q*-Value
Complement system						
-.GVFVLNK.-	C3	Complement component C3	-0.235	<0.001	-3.211	<0.001
K.KVFLDC*C*NYITELRR.Q	C3	Complement component C3	0.575	<0.001	na	na
R.VVLVAVDK.G	C3	Complement component C3	-2.323	<0.001	na	na
K.YFKPGM#PFDLM#VFVTNPDGSPAYR.V	C3	Complement component C3	-0.863	0.094	-1.397	<0.001
R.IPIEDGSGEVVLSR.K	C3	Complement component C3	-2.998	0.094	na	na
K.PGFTIVGPNSVQC*YHFGLSPDLPIC*K.E	CFH	Complement factor H	-0.547	<0.001	na	na
K.SSNLIILEEHLK.N	CFH	Complement factor H	-0.705	0.059	na	na
K.VKDISEVVTPR.F	CFB	Complement factor B	0.603	<0.001	0.603	0.087
R.RPASPISTIQPK.A	C8G	Complement component 8, gamma polypeptide	0.801	<0.001	na	na
R.VPANLENVGFEVQTAEDDLKTDFYK.D	C6	Complement component 6	na	na	-2.450	<0.001
R.HLVPGAPFLLQALVR.E	C4A	Complement component 4A (Rodgers blood group)	na	na	-3.720	0.087
R.LLEPHC*FPLSLVPTEFC*PSPPALK.D	C7	Complement component 7	na	na	-2.111	0.087
Coagulation system						
R.LTIGEGQQHHLGGAK.Q	FGG	Fibrinogen gamma chain	1.145	0.059	na	na
K.EKGEIQNILQK.V	KLKB1	Kallikrein B, plasma (Fletcher factor) 1	2.543	<0.001	na	na
K.FEVQVTVPK.I	A2M	Alpha 2 macroglobulin	-0.329	<0.001	na	na
R.KAAISGENAGLVR.A	ITIH1	Inter-alpha (globulin) inhibitor H1	0.272	<0.001	na	na
K.GFPIKEDFLEQSEQLFGAKPVSLTGK.Q	SERPINF2	Serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 2	na	na	-2.183	0.087
K.TSDQIHFFFAK.L	SERPINC1	Serpin peptidase inhibitor, clade C (antithrombin), member 1	na	na	-4.296	0.087
Antiprotease						
R.NLAVSQVVHK.A	SERPINA3	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	-0.226	<0.001	na	na
K.VLSALQAVQGLLVAQGR.A	AGT	Angiotensinogen	-1.132	<0.001	-1.132	<0.001
K.DPTFIPAPIQAK.T	AGT	Angiotensinogen	-0.319	<0.059	na	na
Metabolic						
R.EYSGTIASEANTYLNSK.S	APOB	Apo-B100 precursor	0.771	0.059	na	na
K.DKDQEVLLQTFLDDASPGDKR.L	APOB	APOB protein	na	na	-2.169	<0.001
R.ILGEELGFASLHDLQLLGK.L	APOB	Apolipoprotein B (including Ag(x) antigen)	na	na	-2.646	0.087
K.KLVPFATELHER.L	APOA4	Apolipoprotein A-IV	-0.458	<0.001	-0.458	<0.001
K.FLNVLSPR.G	IGFBP3	Insulin-like growth factor binding protein 3	na	na	-3.154	<0.001
R.VAGLLEDTFPGLLGLR.V	IGFALS	Insulin-like growth factor binding protein, acid labile subunit	na	na	-2.214	0.087
Other						
R.C*EGPIPDVTFELLR.E	A1BG	Alpha-1-B glycoprotein	-1.037	0.059	-2.196	<0.001
K.NGVAQEPVHLDSPAIK.H	A1BG	Alpha-1-B glycoprotein	0.238	<0.001	0.238	0.087
K.SEDC*FILDHGK.D	GSN	Gelsolin (amyloidosis, Finnish type)	na	na	-0.303	0.087

na, Not applicable. False discovery rate >10% for this analysis method.

[a]A negative regression coefficient estimate indicates decreasing peptide abundance levels across the three study groups, from controls to SLW to RPD, while a positive estimate indicates increasing peptide abundance levels across the three study groups.

mapped to alpha-1B-glycoprotein, a plasma protein of unknown function.

Pathway Studio™ identified nine other proteins with multiple connections to the 21 proteins identified as being potentially differentially expressed across the three groups in this study (Figure 2). These included insulin (INS), plasminogen (PLG), fibrinogen alpha (FGA), coagulation factor 2 (F2), interleukin (IL)-6 and IL-1β, signal transducer and activator of transcription 3 (STAT3), cyclin-dependent kinase 2 (CDK2) and a protein predicted to be similar to RIKEN (FAM3A). Changes in the regulation of plasminogen and its role in coagulation have been associated with smoking and COPD (Tapson 2005). IL-6 and IL-1β promote the inflammatory response and both have been observed to be increased in the sputum (and IL-6 in serum) of smokers and persons with COPD (Chung 2001, Karadag et al. 2008, Walter et al. 2008, Yanbaeva et al. 2007). Furthermore, in a large genome-wide association study, a specific small nucleotide polymorphism found in the *IL6* receptor gene was identified as associated with COPD (Wilk et al. 2007). This suggests that IL-6 signalling may be an important pathway in COPD. A *IL1B* gene polymorphism has also been linked with COPD in a Korean population (Lee et al. 2008). In addition to the human data linking *IL1B* to COPD, a recent mouse model overexpressing *IL1B* in the lung demonstrated similar tissue changes with inflammation, tissue remodelling and distal airway enlargement (Lappalainen et al. 2005). Of the nine additional proteins found by pathway analysis, insulin, plasminogen, IL-6 and IL-1β had the greatest number of interactions with our 21 differentially expressed proteins, suggesting that the these additional proteins may represent common mechanistic pathways for COPD in cigarette smokers and for rate of lung function decline in COPD.

Discussion

Using high-resolution MS proteomics and two rigorous statistical methods, we identified multiple peptides whose expression linearly correlated across three groups

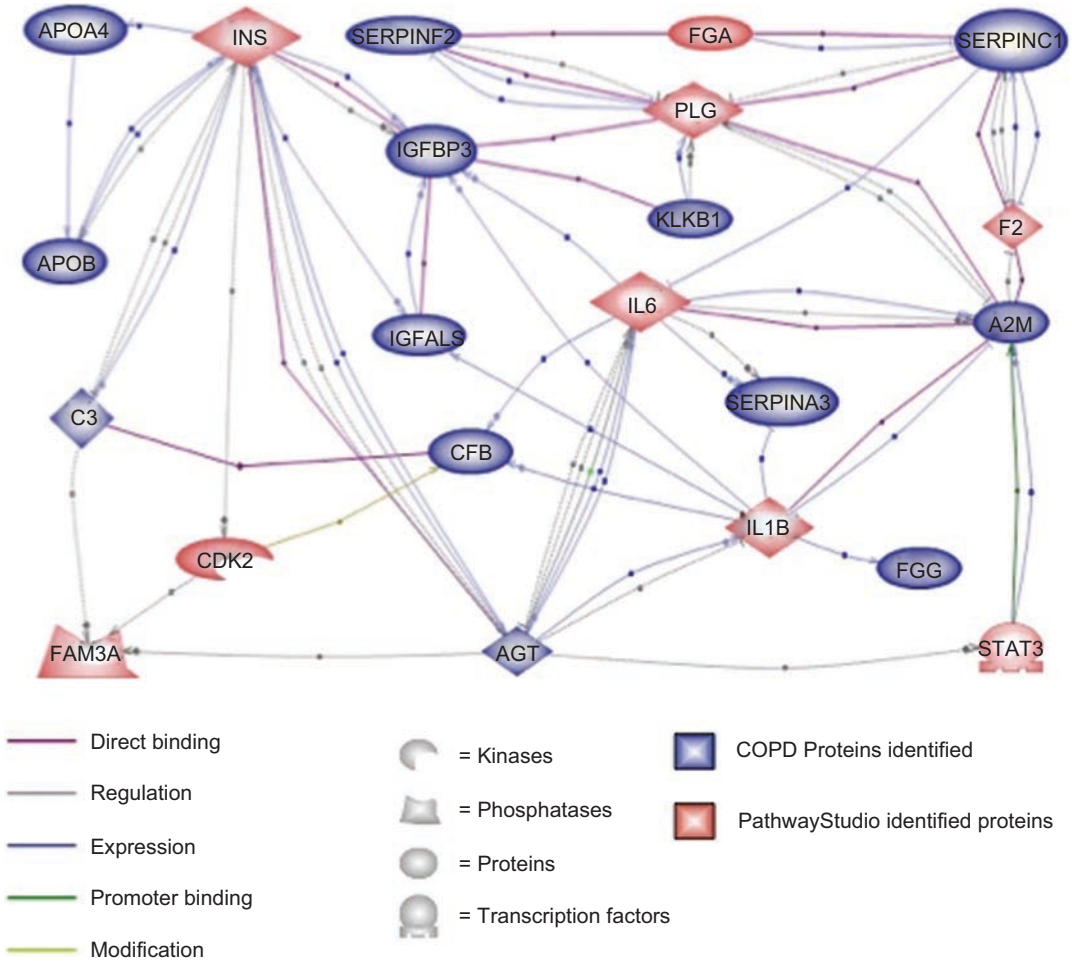

Figure 2. PathwayStudio analysis of direct interactions with other proteins. Proteins identified by PathwayStudio are shown in red; INS, insulin 2 (proinsulin); PLG, plasminogen; FGA, fibrinogen, alpha polypeptide; IL, interleukin; F2, coagulation factor II; STAT3, signal transducer and activator of transcription 3; CDK2, cyclin-dependent kinase 2; FAM3A, similar to RIKEN cDNA 1810037C20 (predicted).

of cigarette smokers classified spirometrically as having COPD with slow or no lung function decline, COPD with rapid decline and an unaffected control group. Thirty unique peptides, representing 21 proteins, differentiated the three groups. The majority of the peptides observed were components of the complement or coagulation cascades, consistent with the chronic and abnormal inflammatory response that is the hallmark of COPD and which is often associated with a prothrombotic state (Tapson 2005). Pathway Studio™ analysis identified nine additional proteins that had multiple interactions with the 21 observed proteins. Interestingly, the four proteins with the greatest number of interactions with our 21 differentially expressed proteins were insulin, plasminogen, IL-6, and IL-1β, all of which have been previously associated with COPD or its complications. Insulin resistance, metabolic syndrome and diabetes have been shown to be associated with COPD (Bolton et al. 2007, Tiengo et al. 2008, Wouters et al. 2002). Both COPD and metabolic syndrome/insulin resistance appear to be systemic proinflammatory, prothrombotic disorders with significant associated, and often common, comorbidities (Bolton et al. 2007, Sevenoaks & Stockley 2006, Tiengo et al. 2008). What is not clear is whether the systemic inflammation is a primary or secondary phenomenon and whether it precedes or follows the airflow decline.

There is increasing evidence that the clinical features of COPD correlate poorly with airflow limitation as measured by spirometry (Fabbri & Rabe 2007) and, therefore, that spirometric parameters alone are inadequate as diagnostic and prognostic biomarkers for this complex disease (Stockley 2007). A more comprehensive evaluation using a multidimensional index (BODE) that incorporates *b*ody mass index, airflow *o*bstruction, *d*yspnea, and *e*xercise capacity, has been shown to be more predictive of mortality than FEV_1 alone (Celli et al. 2004).

Although off-line peptide fractionation enables the identification of a greater number of low-abundance plasma proteins, a disadvantage is the added instrument time required for the data collection from each sample. In this study, each pooled sample was fractionated off-line into ten well-separated fractions, thus increasing the data collection time by a factor of 10. In the interest of reasonable data collection times, the 18 plasma samples in each study group were grouped into three pools of six samples each, for a total of 90 RP-LC-MS/MS samples in the study, each requiring approximately 2 h per data collection (not including blanks and quality controls collected every ten samples). A disadvantage of sample pooling is the inability to collect information on individual variation. However, a benefit is the dilution of undesired individual variation (noise), and the amplification of any signal, by the factor of dilution (i.e. 6 in this study).

This study had some limitations. Categorization of the subjects with COPD by rapid or slow FEV_1 decline was based on each subject's annualized rate of decline as determined from the slope of a linear regression of six FEV_1 measurements over the 5 years of participation in the Lung Health Study. This determination was not adjusted for factors known to impact lung function decline, such as sustained or intermittent smoking cessation (Anthonisen et al. 1994, 2002), frequency of bronchopulmonary infections (Watson et al. 2006) and use of inhaled glucocorticoids (Soriano et al. 2007). It is possible that these factors may also have had differential effects on plasma protein expression, and a limitation of the proteomics analysis is that it did not adjust for these factors. We do note that the proportion of subjects in each group that had ceased smoking by the GAP assessment were not significantly different (see footnote in Table 1). The proteomic analysis was performed on plasma obtained an average of 12 years after the determination of lung function decline category was made, without reassessment of rate of decline during the intervening 12 years. However, low-grade, chronic airway and systemic inflammation in COPD has been shown to persist even years after quitting smoking (Agusti et al. 2003, Agusti & Soriano 2008, Hogg 2006), as well as autoantibody evidence of a self-reactive adaptive immune response (Feghali-Bostwick et al. 2008). Although peripheral plasma proteins may not be the best reflection of intrapulmonary processes resulting in airflow decline, peripheral blood sampling is relatively non-invasive, and COPD is increasingly recognized as a systemic inflammatory disorder with a pulmonary component (Agusti & Soriano 2008, Fabbri et al. 2008).

Further investigation will be required to determine whether the 21 plasma proteins whose expression was differentially associated (linearly) with spirometric lung function and its rate of decline are the cause or the result of the lung function decline. In summary, the 30 plasma peptides mapping to 21 proteins were linearly differentially expressed across cigarette smokers with COPD and rapid FEV_1 decline, COPD with slow or no decline and those without COPD and may provide valuable mechanistic insight into COPD disease progression. The authors recognize that confirmation in an independent sample is the next step in this research and unfortunately the authors were unable to collect these data. Further investigation will also be necessary to determine whether the 21 proteins are the cause or the result of lung function decline.

Acknowledgements

The authors gratefully acknowledge the contributions to this study and manuscript by Michael S. Paul, PhD and Alex Lindell from LineaGen, Inc., Salt Lake City, UT and George J. Patskan, PhD from Altria Client Services. The

authors also acknowledge the comments of reviewers Michael Oldham, PhD and Marc R. Kraus, PhD, the editorial assistance of Eileen Y. Ivasauskas of Accuwrit Inc., data management by Zaigang Liu, and sample preparation by Rebecca Secrist. Financial support was provided by Philip Morris USA Inc. and LineaGen, Inc.

Declaration of interest

E.vdO. was a consultant for Altria Client Services. The Center for Biomarker Research and Personalized Medicine was made possible in part by a start-up gift from Philip Morris USA Inc. to the VCU School of Pharmacy. T.B.L., J.S.E., W.K. and J.W.F. are employees of Altria Client Services. E.L.M. and B.K.Z. were employees of Altria Client Services at the time of manuscript preparation. The other authors declare no potential conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Agusti A, MacNee W, Donaldson K, Cosio M. (2003). Hypothesis: does COPD have an autoimmune component? *Thorax* 58:832–4.

Agusti A, Soriano JB. (2008). COPD as a systemic disease. *COPD* 5:133–8.

Anthonisen NR, Connett JE, Kiley JP, Altose MD, Bailey WC, Buist AS, Conway WA Jr, Enright PL, Kanner RE, O'Hara P et al. (1994). Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV_1. The Lung Health Study. *JAMA* 272:1497–505.

Anthonisen NR, Connett JE, Murray RP. (2002). Smoking and lung function of Lung Health Study participants after 11 years. *Am J Respir Crit Care Med* 166:675–9.

Barnes PJ, Shapiro SD, Pauwels RA. (2003). Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J* 22:672–88.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B* 57:289–300.

Bolger MS, Ross DS, Jiang H, Frank MM, Ghio AJ, Schwartz DA, Wright JR. (2007). Complement levels and activity in the normal and LPS-injured lung. *Am J Physiol Lung Cell Mol Physiol* 292:L748–59.

Bolton CE, Evans M, Ionescu AA, Edwards SM, Morris RH, Dunseath G, Luzio SD, Owens DR, Shale DJ. (2007). Insulin resistance and inflammation – a further systemic complication of COPD. *COPD* 4:121–6.

Carpenter SL, Mathew P. (2008). Alpha2-antiplasmin and its deficiency: fibrinolysis out of balance. *Haemophilia* 14:1250–4.

Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, Pinto Plata V, Cabral HJ. (2004). The body-mass index airflow obstruction dyspnea and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 350:1005–12.

Choolani M, Narasimhan K, Kolla V, Hahn S. (2009). Proteomic technologies for prenatal diagnostics: advances and challenges ahead. *Expert Rev Proteomics* 6:87–101.

Chung KF. (2001). Cytokines in chronic obstructive pulmonary disease. *Eur Respir J* (Suppl. 34):50s–59s.

Elias JE, Gygi SP. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* 4(3):207–214.

Ezzati M, Lopez AD. (2003). Estimates of global mortality attributable to smoking in 2000. *Lancet* 362:847–52.

Fabbri LM, Luppi F, Beghe B, Rabe KF. (2008). Complex chronic comorbidities of COPD. *Eur Respir J* 31:204–12.

Fabbri LM, Rabe KF. (2007). From COPD to chronic systemic inflammatory syndrome? *Lancet* 370:797–9.

Fay WP, Garg N, Sunkar M. (2007). Vascular functions of the plasminogen activation system. *Arterioscler Thromb Vasc Biol* 27:1231–7.

Feghali-Bostwick CA, Gadgil AS, Otterbein LE, Pilewski JM, Stoner MW, Csizmadia E, Zhang Y, Sciurba FC, Duncan SR. (2008). Autoantibodies in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 177:156–63.

Fu Q, Van Eyk JE. (2006). Proteomics and heart disease: identifying biomarkers of clinical utility. *Expert Rev Proteomics* 3:237–49.

Furie B, & Furie BC. (2005). Thrombus formation *in vivo. J Clin Invest* 115:3355–62.

Hersh CP, DeMeo DL, Silverman EK. (2008). National Emphysema Treatment Trial state of the art: genetics of emphysema. *Proc Am Thorac Soc* 5:486–93.

Hogg JC. (2006). Why does airway inflammation persist after the smoking stops? *Thorax* 61:96–7.

Hou Y, Okada K, Okamoto C, Ueshima S, Matsuo O. (2008). Alpha2-antiplasmin is a critical regulator of angiotensin II-mediated vascular remodeling. *Arterioscler Thromb Vasc Biol* 28(7):1257–62.

Ishii T, Matsuse T, Teramoto S, Matsui H, Hosoi T, Fukuchi Y, Ouchi Y. (2000). Association between alpha-1-antichymotrypsin polymorphism and susceptibility to chronic obstructive pulmonary disease. *Eur J Clin Invest* 30:543–8.

Jacot W, Lhermitte L, Dossat N, Pujol JL, Molinari N, Daures JP, Maudelonde T, Mange A, Solassol J. (2008). Serum proteomic profiling of lung cancer in high-risk groups and determination of clinical outcomes. *J Thorac Oncol* 3:840–50.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36(database issue):D480–4.

Kanehisa M, Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30.

Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34(database issue):D354–7.

Karadag F, Karul AB, Cildag O, Yilmaz M, Ozcan H. (2008). Biomarkers of systemic inflammation in stable and exacerbation phases of COPD. *Lung* 186:403–9.

Kumar R, Singh VP, Baker KM. (2008). The intracellular renin-angiotensin system: implications in cardiovascular remodeling. *Curr Opin Nephrol Hypertens* 17:168–73.

Lalouel JM, Rohrwasser A, Terreros D, Morgan T, Ward K. (2001). Angiotensinogen in essential hypertension: from genetics to nephrology. *J Am Soc Nephrol* 12:606–15.

Lappalainen U, Whitsett JA, Wert SE, Tichelaar JW, Bry K. (2005). Interleukin-1beta causes pulmonary inflammation emphysema and airway remodeling in the adult murine lung. *Am J Respir Cell Mol Biol* 32:311–18.

Lee JM, Kang YR, Park SH, Cha SI, Kim JS, Kang HK, Lee WK, Kim MJ, Kim CH, Kim NS, Jung TH, Park JY. (2008). Polymorphisms in interleukin-1B and its receptor antagonist genes and the risk of chronic obstructive pulmonary disease in a Korean population: a case–control study. *Respir Med* 102:1311–20.

Lee PS, Waxman AB, Cotich KL, Chung SW, Perrella MA, Stossel TP. (2007). Plasma gelsolin is a marker and therapeutic agent in animal sepsis. *Crit Care Med* 35:849–55.

Li X, Gong Y, Wang Y, Wu S, Cai Y, He P, Lu Z, Ying W, Zhang Y, Jiao L, He H, Zhang Z, He F, Zhao X, Qian X. (2005). Comparison of alternative analytical techniques for the characterisation of the human serum proteome in HUPO Plasma Proteome Project. *Proteomics* 5:3423–41.

Mannino DM, Homa DM, Akinbami LJ, Ford ES, Redd SC. (2002). Chronic obstructive pulmonary disease surveillance – United States 1971–2000. *MMWR Surveill Summ* 51:1–16.

Marcovina S, Packard CJ. (2006). Measurement and meaning of apolipoprotein AI and apolipoprotein B plasma levels. *J Intern Med* 259:437–46.

Markiewski MM, Lambris JD. (2007). The role of complement in inflammatory diseases from behind the scenes into the spotlight. *Am J Pathol* 171:715–27.

Meng Z, Veenstra TD. (2007). Proteomic analysis of serum plasma and lymph for the identification of biomarkers. *Proteomics Clin Appl* 1:747–57.

Navabi A, Mather DE, Bernier J, Spaner DM, Atlin GN. (2009). QTL detection with bidirectional and unidirectional selective genotyping: marker-based and trait-based analyses. *Theor Appl Genet* 118:347–58.

Omenn GS, States DJ, Adamski M, Blackwell TW, Menon R, Hermjakob H, Apweiler R, Haab BB, Simpson RJ, Eddes JS, et al. (2005). Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 5:3226–245.

Pinto-Plata V, Toso J, Lee K, Park D, Bilello J, Mullerova H, De Souza MM, Vessey R, Celli B. (2007). Profiling serum biomarkers in patients with COPD: associations with clinical parameters. *Thorax* 62:595–601.

Poller W, Faber JP, Scholz S, Weidinger S, Bartholome K, Olek K, Eriksson S. (1992). Mis-sense mutation of alpha 1-antichymotrypsin gene associated with chronic lung disease. *Lancet* 339:1538.

Qin X, Swertfeger DK, Zheng S, Hui DY, Tso P. (1998). Apolipoprotein AIV: a potent endogenous inhibitor of lipid oxidation. *Am J Physiol* 274:H1836–40.

Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, Zielinski J. (2007). Global strategy for the diagnosis management and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 176:532–55.

Rahman I, Morrison D, Donaldson K, MacNee W. (1996). Systemic oxidative stress in asthma COPD and smokers. *Am J Respir Crit Care Med* 154:1055–60.

Ribeiro-Oliveira A, Jr, Nogueira AI, Pereira RM, Boas WW, Dos Santos RA, Simoes e Silva AC. (2008). The renin–angiotensin system and diabetes: an update. *Vasc Health Risk Manag* 4:787–803.

Rosenberg RD, Bauer KA. (1987). Thrombosis in inherited deficiencies of antithrombin protein C and protein S. *Hum Pathol* 18:253.

Rosenzweig CN, Zhang Z, Sun X, Sokoll LJ, Osborne K, Partin AW, Chan DW. (2009). Predicting prostate cancer biochemical recurrence using a panel of serum proteomic biomarkers. *J Urol* 181:1407–14.

Sevenoaks M, J, & Stockley R. A. (2006). Chronic obstructive pulmonary disease inflammation and co-morbidity – a common inflammatory phenotype? *Respir Res* 7:70.

Soriano JB, Sin DD, Zhang X, Camp PG, Anderson JA, Anthonisen NR, Buist AS, Burge PS, Calverley PM, Connett JE, Petersson S, Postma DS, Szafranski W, Vestbo J. (2007). A pooled analysis of FEV1 decline in COPD patients randomized to inhaled corticosteroids or placebo. *Chest* 131:682–9.

Stockley RA. (2007). Biomarkers in COPD: time for a deep breath. *Thorax* 62:657–60.

Storey JD, Tibshirani R. (2003a). Statistical methods for identifying differentially expressed genes in DNA microarrays. *Methods Mol Biol* 224:149–57.

Storey JD, Tibshirani R. (2003b). Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440–5.

Tapson VF. (2005). The role of smoking in coagulation and thromboembolism in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2:71–7.

Tiengo A, Fadini GP, Avogaro A. (2008). The metabolic syndrome diabetes and lung dysfunction. *Diabetes Metab* 34:447–54.

van den Oord EJ. (2005). Controlling false discoveries in candidate gene studies. *Mol Psychiatry* 10:230–1.

Vivanco F, Martin-Ventura JL, Duran MC, Barderas MG, Blanco-Colio L, Darde VM, Mas S, Meilhac O, Michel JB, Tunon J, Egido J. (2005). Quest for novel cardiovascular biomarkers by proteomic analysis. *J Proteome Res* 4:1181–91.

Voelkel NF, Cool CD. (2003). Pulmonary vascular involvement in chronic obstructive pulmonary disease. *Eur Respir J* (Suppl. 46):28s–32s.

Voelkel NF, Douglas IS, Nicolls M. (2007). Angiogenesis in chronic lung disease. *Chest* 131:874–9.

Walter RE, Wilk JB, Larson MG, Vasan RS, Keaney JF Jr, Lipinska I, O'Connor GT, Benjamin EJ. (2008). Systemic inflammation and COPD: the Framingham Heart Study. *Chest* 133:19–25.

Watson L, Vonk JM, Lofdahl CG, Pride NB, Pauwels RA, Laitinen LA, Schouten JP, Postma DS. (2006). Predictors of lung function and its decline in mild to moderate COPD in association with gender: results from the Euroscop study. *Respir Med* 100:746–53.

Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT. (2007). Framingham Heart Study genome-wide association: results for pulmonary function measures. *BMC Med Genet* 8 (Suppl. 1):*S8.*

Wouters EF, Creutzberg EC, Schols AM. (2002). Systemic effects in COPD. *Chest* 121 (Suppl.):127S–30S.

Yanbaeva DG, Dentener MA, Creutzberg EC, Wesseling G, Wouters EF. (2007). Systemic effects of smoking. *Chest* 131:1557–66.