

Gender differences in the bronchoalveolar lavage cell proteome of patients with chronic obstructive pulmonary disease

Maxie Kohler, PhD,^a AnnSofi Sandberg, MSc,^a Sanela Kjellqvist, PhD,^{b*} Andreas Thomas, PhD,^c Reza Karimi, MD,^a Sven Nyrén, MD, PhD,^a Anders Eklund, MD, PhD,^a Mario Thevis, PhD,^c C. Magnus Sköld, MD, PhD,^a and Åsa M. Wheelock, PhD^a *Stockholm, Sweden, and Cologne, Germany*

Background: Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity and mortality worldwide and is increasing, primarily among women. Underdiagnosis is common, and because of the heterogeneous disease characteristics, molecular markers of specific disease phenotypes and more efficacious treatment regimens are urgently needed.

Objective: In this study the soluble proteome of bronchoalveolar lavage cells, primarily consisting of macrophages, was investigated with the aim of identifying phenotypic differences in early disease development.

Methods: Two-dimensional difference gel electrophoresis was used for relative quantification of protein levels, and multivariate modeling was applied to identify proteins of interest that were subsequently identified by means of liquid chromatography–mass spectrometry.

Results: Significant gender differences were unveiled, with numerous alterations in the bronchoalveolar lavage cell proteome occurring in female but not male patients with COPD. Specifically, a subset of 19 proteins provided classification of female healthy smokers from female patients with COPD with 78% predictive power. Subsequent pathway analyses linked the observed alterations to downregulation of the lysosomal

pathway and upregulation of the oxidative phosphorylation pathway, possibly linking dysregulation of macroautophagy to a female-dominated COPD disease phenotype.

Conclusion: This investigation makes an important contribution to the elucidation of putative molecular mechanisms underlying gender-based differences in the pathophysiology of COPD, linking alterations of specific molecular pathways to previously observed gender differences in clinical COPD phenotypes. Furthermore, these results stress the importance of the gender-specific search for biomarkers, diagnosis, and treatment in COPD. (*J Allergy Clin Immunol* 2012;■■■:■■■-■■■.)

Key words: Chronic obstructive pulmonary disease, alveolar macrophage, proteomics, multivariate modeling, gender, difference gel electrophoresis, bronchoalveolar lavage

Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity and mortality worldwide^{1,2} and is characterized by progressive airflow limitations that are not fully reversible. The main risk factor for development of COPD is smoking, with 15% to 44% of smokers developing disease.^{3,4} COPD is a heterogeneous disease⁵ potentially representing several disease phenotypes with different clinical presentations, outcomes, risk factors, and causes. After being predominantly a male-associated disease for decades, the mortality rate of COPD in women now exceeds that of men.⁶ The disease phenotype appears to be different between men and women,^{3,7} with the male-associated disease phenotype characterized by a higher degree of emphysema in relation to the cumulative smoking history.⁸⁻¹¹ It has been suggested that smoking has a more detrimental effect on lung function in women,^{3,12,13} particularly after menopause.¹² Taken together, this indicates that disease development might have different courses in men and women, with divergent mechanisms that might benefit from different diagnostic and treatment strategies.

Proteomics studies have previously been used for biomarker discovery in patients with COPD.¹⁴⁻²⁶ As an example, Plymoth et al²⁴ identified a set of prognostic markers that discriminated male patients with COPD from healthy subjects 6 to 7 years before the manifestation of clinical symptoms by multivariate modeling. Verrills et al¹⁸ identified a set of 4 potential plasma markers (α_2 -macroglobulin, haptoglobin, ceruloplasmin, and hemopexin) to differentiate between asthmatic patients, patients with COPD (Global Initiative for Chronic Obstructive Lung Disease [GOLD] stage II-IV), and healthy control subjects. In spite of the apparent gender differences in susceptibility and disease course, very few studies are reporting gender-specific analyses. De Torres et al²⁶ reported significant gender differences in a subset of serum markers (IL-6, IL-16, and vascular endothelial

From ^athe Respiratory Medicine Unit, Department of Medicine Solna, Center for Molecular Medicine, Karolinska Institutet/University Hospital, Stockholm; ^bthe Atherosclerosis Research Unit, Center for Molecular Medicine, Karolinska University Hospital, Stockholm; and ^cthe Center for Preventive Doping Research/Institute of Biochemistry, German Sport University Cologne.

*Previously Sanela Kurtovic.

Supported by the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research (SSF), VINNOVA (VINN-MER), EU Fp6 Marie Curie, the Karolinska Institutet, AFA Insurances for financial support, the King Oscar II Jubilee Foundation, the Mats Kleberg Foundation, King Gustaf V's, and Queen Victoria's Freemasons, and Hesselman's Foundations, the regional agreement on medical training and clinical research (ALF) between Stockholm County Council, and the Karolinska Institutet.

Disclosure of potential conflict of interest: M. Kohler has received grants from the Karolinska Institutet and AFA Insurances. S. Kjellqvist has received grants from the Swedish Research Council. A. Eklund has received grants from the Swedish Heart Lung Foundation. A. M. Wheelock has received grants from the Swedish Heart Lung Foundation, the Swedish Foundation for Strategic Research, VINNOVA, EU Fp6 Marie Curie, and the King Oscar II Jubilee Foundation. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication May 25, 2012; revised August 20, 2012; accepted for publication September 24, 2012.

Corresponding author: Åsa M. Wheelock, PhD, Lung Research Lab L4:01, Respiratory Medicine Unit, Department of Medicine, Karolinska Institutet, 17176 Stockholm, Sweden. E-mail: asa.wheelock@ki.se. Or: Maxie Kohler, PhD, Physical Chemistry Department, University of Cologne, Luxemburger Str 116, 50939 Cologne, Germany. E-mail: maxie.kohler@uni-koeln.de.

0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology
http://dx.doi.org/10.1016/j.jaci.2012.09.024

Abbreviations used

BAL:	Bronchoalveolar lavage
COPD:	Chronic obstructive pulmonary disease
CV-ANOVA:	Cross-validated ANOVA
2D-DIGE:	Two-dimensional difference gel electrophoresis
2DE:	Two-dimensional electrophoresis
FDR:	False discovery rate
GOLD:	Global Initiative for Chronic Obstructive Lung Disease
OPLS:	Orthogonal projections to latent structures
PCA:	Principal component analysis
PLS:	Partial least squares projections to latent structures
SUS:	Shared and unique structures
VIP:	Variable Influence on the Projections

growth factor). Larsson et al²⁷ reported a significantly higher blood neutrophil concentration in female compared with male patients with COPD with severe disease.

In the current study we used 2-dimensional difference gel electrophoresis (2D-DIGE) to investigate the soluble proteome of immune cells (mainly macrophages) obtained by means of bronchoalveolar lavage (BAL), with special focus on the differentiation between the genders in early disease stages. Macrophages from patients with COPD have been shown to secrete more inflammatory and elastolytic proteins^{28,29} compared with those from healthy smokers and are localized at the sites of alveolar destruction, and differences in alveolar macrophage polarization between patients with COPD and healthy smokers have been reported.^{30,31} Taken together, these facts make them a relevant target with respect to biomarker discovery, as well as for investigation of the pathophysiology of COPD.

METHODS

Detailed descriptions of the methods used in this study are provided in the **Methods** section in this article's Online Repository at www.jacionline.org.

Subjects and study design

The study was carried out on subjects from the Karolinska COSMIC cohort at the Karolinska University Hospital Solna, Karolinska Institutet, Sweden, encompassing 120 subjects matched in terms of age (45–65 years) and gender (20/20 per group) from groups of healthy never smokers, smokers with normal lung function, and patients with COPD with mild-to-moderate disease (GOLD stage I–II; FEV₁, 50% to 100%; and FEV₁/forced vital capacity ratio, <0.7). Volunteers did not use inhaled or oral corticosteroids and had no exacerbations for at least 3 months before inclusion in the study. All subjects underwent clinical examination, chest radiography, computed tomography, and spirometry. Lung function parameters were calculated as postbronchodilator percent predicted values by using the European Community of Coal and Steel normal values.³² Patients with COPD and healthy smokers were matched in terms of smoking history (>10 pack years, >10 cigarettes/d in the past 6 months). The COPD group consisted of both current smokers and exsmokers (≥2 years since stopping smoking). The study was approved by the Stockholm Regional Ethical Board (no. 2006/959-31/1), and participants provided written consent.

Sample collection and 2D-DIGE

Samples were collected by means of BAL, as previously described.^{33,34} After lysis of cells in 8 mol/L urea, 2 mol/L thiourea, 4% Chaps, and 33 mmol/L Tris, aliquots of 50-μg samples were labeled with minimal DIGE, according to the supplier's recommendations. A triplex of 2 samples and

1 internal standard were coseparated by using isoelectric focusing with 18-cm strips, pH 4 to 7, for 86 kVhrs and SDS-PAGE was performed on lab-casted 10% Tris-glycine gels before image acquisition with an FLA Typhoon 9000 laser scanner (GE Healthcare, Uppsala, Sweden).

Quantitation and statistical analyses

Image analysis and univariate statistics were performed with SameSpots version 4.0 (Nonlinear Dynamics, Newcastle, United Kingdom), *P* values were calculated by using ANOVA, and *q* values represent *P* values corrected for false discovery rate (FDR).³⁵

Multivariate statistical modeling was performed with SIMCA P+ version 13.0 (Umetrics, Umeå, Sweden), applying both unsupervised principal component analysis (PCA)³⁶ and supervised orthogonal projections to latent structures (OPLS) analysis.³⁷ Before analysis, data were log transformed, normalized, scaled to unit variance, and mean centered to avoid weighting based on protein abundance in the analyses. Model performance is reported as cumulative correlation coefficients for the model (*R*²) and predictive performance based on 7-fold cross-validation calculations (*Q*²), as well as cross-validated ANOVA (CV-ANOVA) *P* values for the OPLS models. Variable selection for model optimization was performed iteratively based on the Variable Influence on the Projections (VIP) and scaled loadings of the predictive component of the OPLS models (p[corr]).

PCA is a powerful multivariate tool commonly used for quality control and outlier identification, as well as to investigate the overall trends in a dataset. In contrast to the more commonly used PCA modeling, OPLS analysis is a supervised method designed to separate structured noise unrelated (orthogonal) to the predictive variance of interest (eg, between patients with COPD and smokers), thereby increasing the interpretability of the multivariate model, particularly in deriving the observed group separation back to the protein variables of interest. The results are displayed in a score plot displaying the separation of groups (predictive component) along the y-axis and within-group variation between subjects (orthogonal components) in the x-axis, as well as the corresponding loadings plot or variable importance plot, showing the variables' relative importance for separating the groups.

Protein identification by means of mass spectrometry and pathway analysis

Preparative gels (150–250 μg of protein) were stained with silver according to the method of Yan et al.³⁸ Gel plugs of selected proteins were subjected to in-gel trypsin digestion, peptides were analyzed by means of nano-LC-Orbitrap-MS, and protein identification was performed with Proteome Discoverer 1.0 using Sequest,³⁹ searching the human FASTA database (2006). Pathway analysis was performed with the DAVID tool^{40,42} using KEGG pathways.^{43,44}

Validation by means of Western blotting

Western blotting was performed for 2 proteins differentially expressed between smokers and patients with COPD: lamin A/C and cathepsin B (7 smokers and 7 patients with COPD). The samples nearest to the group means of the relevant 2D-DIGE gel spot volumes (lamin A/C, spot 1566; cathepsin B, spot 1113) were selected to ensure representative validation, and band intensities were normalized to actin. Proteins were separated by using SDS-PAGE and blotted to nitrocellulose membranes, and proteins were visualized by using primary anti-cathepsin B and anti-lamin A/C and anti-actin antibodies and secondary CyDye-coupled antibodies.

RESULTS

Statistical analyses were performed by comparing smoking and nonsmoking subjects separately to limit the confounding effects of smoking. Hence the group comparisons were made between healthy smokers versus actively smoking patients with COPD and healthy nonsmoking control subjects versus exsmokers with COPD. In addition, we compared intergroup and intragroup

gender differences. Additional exclusion criteria specifically designed for the proteome analyses were applied to the Karolinska COSMIC cohort. Low recovery of BAL fluid might result in a more proximal sampling of the airways, resulting in a different BAL cell population.⁴⁵ As such, subjects with a recovery of less than 40% were excluded, representing a slightly more stringent exclusion criteria than previously suggested.⁴⁶ The chosen cutoff was validated by means of multivariate statistical modeling, revealing slightly different proteome characteristics of the 2 subjects with 35% to 40% recoveries. At the 40% cutoff, no significant correlation between BAL recoveries and proteome profiles of interest were detected ($R^2 < 0.30$, partial least squares projections to latent structures [PLS] inner relation). Samples with a macrophage percentage composition of less than 85% were excluded to ensure sufficient enrichment of the target cell population (ie, alveolar macrophages). Table I shows the groups and the clinical characteristics of the subjects included in the study. The highly reproducible DIGE multiplexing approach resulted in average within-group coefficients of variance of 22% to 42% for all 8 groups outlined in Table I, resulting in a statistical power of greater than 93% for all comparisons of interest. These values are comparable with our previous in-depth characterizations of the sources of variance for the work flow⁴⁷ and, because of the correction for technical variance by means of an internal standard, primarily represent the biological variance.

Clinical parameters

As determined by using the Student *t* test, the relevant parameters of age, body mass index, macrophage percentage in BAL fluid, and smoking history were not significantly different between any of the groups. The BAL fluid recovery was significantly different between female smokers and female smokers with COPD ($P = .002$), smokers and patients with COPD of both genders ($P = .003$), and healthy volunteers and exsmokers with COPD ($P = .006$). The neutrophil percentage only differed between healthy nonsmokers and exsmokers with COPD ($P = .0002$). As expected, significant differences were also found in the parameters used as inclusion criteria (see Table E1 in this article's Online Repository at www.jacionline.org).

Healthy smokers versus smokers with COPD and gender differences

Univariate analysis revealed 43 protein spots to be differentially expressed between healthy smokers and smokers with COPD ($P < .05$). After *P* value correction to account for multiple testing, 14 of the 43 protein spots were found to be significantly altered ($q < .05$). On analysis by gender, 98 proteins were significantly differentially expressed in women (59 with $q < .05$), and only 6 proteins were significantly altered in the male cohort (none with $q < .05$).

OPLS multivariate modeling was used to analyze differences between healthy smokers and smokers with COPD. OPLS is a supervised method designed to separate structured noise unrelated (orthogonal) to the predictive variance of interest (ie, difference between patients with COPD and smokers), thereby increasing the interpretability of the multivariate model, particularly in deriving the observed group separation back to the protein variables of interest.⁴⁸

After iterative variable selection based on the VIP score, as well as the scaled loadings of the predictive component of the OPLS models ($p[\text{corr}]$), a subset of 17 variables was identified as the most important markers for differentiating between the 2 groups ($R^2 = 0.50$, $Q^2 = 0.45$, $P[\text{CV-ANOVA}] = .0003$; Fig 1, A). Further modeling of men and women separately revealed that the group separation was primarily driven by alterations in the female subjects. The 19 selected variables yielding an optimal separation of the groups in women (Fig 1, B) resulted in a highly significant model ($P[\text{CV-ANOVA}] = 1.1 \times 10^{-7}$) with 78% predictive power ($Q^2 = 0.78$). Correlation analysis with the relative BAL cell composition revealed a correlation between the proteome profile of these subjects and the macrophage percentage composition ($R^2 = 0.51$, PLS inner relation) but not with any other immune cell populations (neutrophils, lymphocytes, basophils, and eosinophils: $R^2 = 0.21$ - 0.23 , PLS inner relation). No significant OPLS model could be achieved for the corresponding male subjects (Fig 1, C). Model statistics for the abovementioned models, as well as protein IDs, are presented in Table II and Table E2 in this article's Online Repository at www.jacionline.org.

Overall, there was very little overlap between the most important proteins in men and women. These differences were further investigated through a shared and unique structures (SUS) plot (Fig 2), in which the scaled loadings (variables/protein spots) of the predictive components from the gender-based models are plotted against each other. The resulting cloud of variables primarily clustering around the y-axis is a result of the much stronger model for the female gender but also indicates that the protein expression pattern in patients with early-stage COPD shares very few common denominators between the 2 genders (high similarity of the models would have resulted in a clustering of variables around the diagonal).

An overview of the different OPLS models and the variables of importance for differentiation between groups is shown in Table II. The variables important for separating female smokers from female patients with COPD are prominent in the gender-mixed comparison as well, as a result of the weak male model. The gender differences are also highlighted when comparing female and male patients with COPD directly (Fig 1, D). Correlation analysis with computed tomographic attenuation values (percentage of total voxels < -900 Hounsfield units) did not reveal any correlation with the gender-related proteome profiles ($R^2 = 0.15$, PLS inner relation). Univariate analysis unmasked 38 spots significantly different between the genders (none with $q < .05$). In contrast, no differences were found between male and female smokers or male and female healthy volunteers by means of univariate or multivariate analysis. A representative 2-dimensional electrophoresis (2DE) gel image annotated with the position of spots/proteins listed in Table II is shown in Fig E1 in this article's Online Repository at www.jacionline.org.

As shown in Table I, all female patients with COPD had undergone menopause, although only 9 of 17 female smokers were postmenopausal. Premenopausal women were excluded in an additional OPLS analysis to investigate the influence of menopause. The resulting model was nearly identical to that including premenopausal women (Fig 2), and no substantial differences in variable contribution were found through SUS plot analysis (see Fig E2 in this article's Online Repository at www.jacionline.org).

TABLE I. Clinical parameters of subjects included in statistical analyses

	Healthy nonsmoker		"Healthy" smoker		Smoker COPD		Exsmoker COPD	
	Male	Female	Male	Female	Male	Female	Male	Female
Group size (no.)	10	13	16	17	8	7	3	3
GOLD (I/II)	NA	NA	NA	NA	5/3	2/5	2/1	2/1
Age (y)	56 ± 9	55 ± 7	54 ± 6	54 ± 7	56 ± 7	59 ± 4	61 ± 7	60 ± 6
BMI (kg/m ²)	27 ± 7	26 ± 4	25 ± 4	25 ± 3	25 ± 4	25 ± 4	26 ± 4	26 ± 3
Recovered BAL (%)	71 ± 7	68 ± 10	55 ± 6	64 ± 9	51 ± 7	51 ± 7	47 ± 6	60 ± 7
Macrophages (%)	92 ± 4	91 ± 4	96 ± 2	96 ± 2	97 ± 2	96 ± 2	91 ± 4	89 ± 1
Neutrophils (%)	1.1 ± 0.6	1.6 ± 1.2	0.9 ± 0.9	0.8 ± 0.8	0.6 ± 0.4	0.6 ± 0.4	2.7 ± 2.2	1.3 ± 0.8
Pack years	0	0	35 ± 15	34 ± 12	42 ± 14	38 ± 10	24 ± 5	24 ± 9
Menopause	NA	10/13	NA	9/17	NA	7/7	NA	3/3
FEV ₁ (%)	116 ± 13	121 ± 13	108 ± 11	110 ± 13	74 ± 10	86 ± 9	85 ± 12	82 ± 5
FEV ₁ /VC ratio (%)	78 ± 7	82 ± 6	75 ± 4	76 ± 6	59 ± 12	62 ± 8	61 ± 10	58 ± 8
FEV ₁ /FVC ratio (%)	81 ± 6	82 ± 6	77 ± 4	78 ± 4	60 ± 7	60 ± 7	65 ± 5	61 ± 5

BMI, Body mass index.

Healthy nonsmokers versus exsmokers with COPD

Additional analyses comparing healthy volunteers and exsmokers with COPD were performed to investigate whether the observed differences are connected to COPD independently from smoking. Applying univariate statistics, 24 spots were significantly differentially expressed ($P < .05$). An OPLS model comparing healthy volunteers and exsmokers with COPD was correlated to the differences found between the corresponding smoking populations through SUS plot analysis, correlating the proteins that were the most important to differentiate between healthy subjects and patients with COPD in smokers (Fig 3, x-axis) versus nonsmokers (Fig 3, y-axis). Overall, the majority of the proteins are closely clustered around the diagonal of the SUS plot, thus validating that the proteome alterations caused by COPD are the same in smoking and nonsmoking subjects. Variables that were less relevant for exsmokers than smokers with COPD were protein spot 432 (vesicle transport-related protein isoform b), 544 (chaperonin), and 504 (no protein identified).

Pathway analysis and protein ontology

Given that the most pronounced differences between healthy smokers and smokers with COPD were observed in the female population, subsequent pathway analyses were performed on data from these groups. Specifically, proteins with VIP > 1.0 (148 proteins) were included in the analysis. Two pathways were found to be significantly altered. The lysosome activity pathway was significantly downregulated in patients with COPD ($P = 3.1 \times 10^{-6}$, 9 altered proteins), with levels of cathepsins (B, D, H, and Z), tripeptidyl peptidase, napsin a, and the glucosidases galactosidase, glucosidase, and hexosaminase B decreased in patients with COPD. In contrast, the oxidative phosphorylation pathway was upregulated in patients with COPD ($P = 5.4 \times 10^{-4}$, 7 altered proteins), primarily because levels of ATP synthases, ATPases, and ubiquinol-cytochrome C reductase were increased in patients with COPD.

In addition to pathway analysis, proteins from Fig 1, B and D, were evaluated and further grouped by gene ontology terms of UniProt (for more information on identified proteins, see Tables E2 and E3 in this article's Online Repository at www.jacionline.org).^{49,50} The most prominent systematic change in female smokers versus female patients with COPD was the upregulated proteins involved in energy metabolism (ATP synthases,

ATPase, and acid α -glucosidase). Furthermore, proteins involved in protein turnover (synthesis/cleavage/transport; spots 1561, 1066, and 1135) were increased. In contrast, expression of the lysosomal protease cathepsin B, as well as structural proteins, such as vimentin or lamin A/C, was downregulated.

Several of the proteins differentially regulated between female smokers and female patients with COPD drove the difference between male and female patients as well, including proteins involved in energy metabolism (ATP synthase and ATPase, spots 447 and 684) and structural components (vimentin, spots 1594 and 1394). Further differences between male and female patients with COPD include tubulin β (spot 1594, less in female patients) or leucine amino peptidase (spot 1305, less in female patients with COPD).

Validation by means of Western blotting

Two of the top-ranked proteins for differentiation between smokers and patients with COPD (spot 1566, lamin A/C; spot 1113, cathepsin B) were analyzed by means of Western blotting to validate the results obtained by using 2D-DIGE. For both proteins, 2 bands were visible on the blots (see Fig E3 in this article's Online Repository at www.jacionline.org) which is in agreement with the multiple protein spots detected by means of 2DE (lamin A/C, spots 392, 395, 397, 1313, 1587, 1600, and 1619; cathepsin B, spots 1026 and 1098). Both bands of cathepsin B (Fig 4, A, and see Fig E3, A, in this article's Online Repository at www.jacionline.org) and lamin A/C (Fig 4, B, and see Fig E3, B) were significantly altered also in the Western blot validation (fold change band 1: 2.5x and band 2: 2.4x for lamin A/C; band 1: 1.6x and band 2: 2.6x for cathepsin B). For cathepsin B, one outlier was excluded from the statistical analysis based on the Dixon Q-test.

DISCUSSION

COPD is increasing in incidence and mortality worldwide, an increase that is primarily driven by the female population. In spite of its prevalence, neither efficacious treatments nor reliable diagnostic tools are available to date.¹ In part, this might be due to COPD being an umbrella diagnosis, with a number of various clinical manifestations resulting in different disease phenotypes.⁵ The heterogeneous disease characteristics of COPD thus require the discovery of novel molecular markers corresponding to various subphenotypes to allow more differentiated and specific

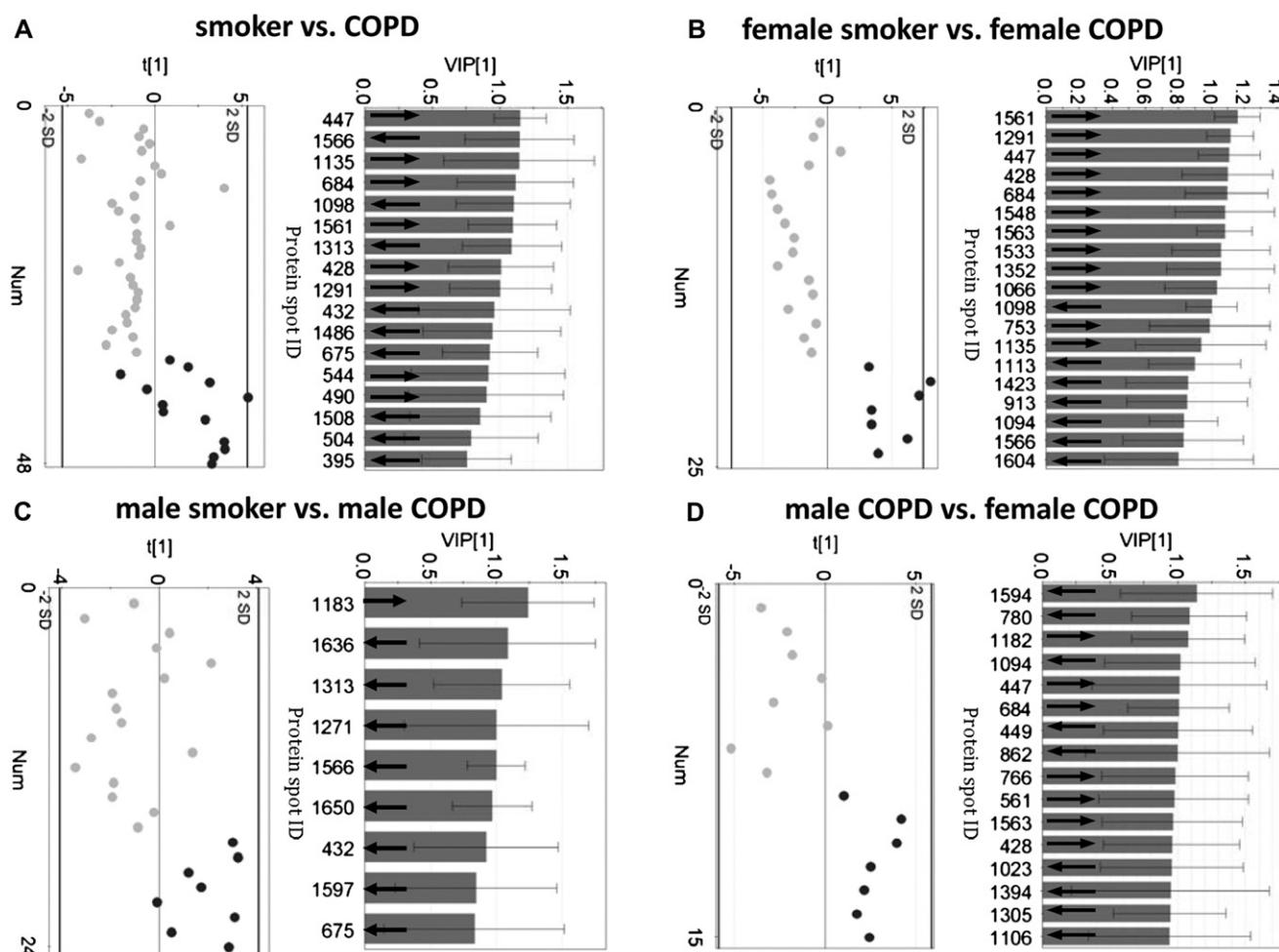


FIG 1. Optimized models showing markers for smokers versus patients with COPD ($R^2 = 0.50$, $Q^2 = 0.45$; **A**), female smokers versus female patients with COPD ($R^2 = 0.81$, $Q^2 = 0.78$; **B**), male smokers versus male patients with COPD ($R^2 = 0.49$, $Q^2 = 0.42$; **C**), and male patients with COPD versus female patients with COPD ($R^2 = 0.76$, $Q^2 = 0.71$; **D**). *Left panels* show score plots (Fig 1, A-C: gray, smokers; black, patients with COPD; Fig 1, D: gray, male patients; black, female patients), with the predictive component along the y-axes. Because no orthogonal components were required, the x-axes merely represent a numeric ordering (Num) of the samples. *Right panels* display the VIP. Arrows indicate upregulation or downregulation in patients with COPD compared with that seen in healthy smokers (Fig 1, A-C) or regulation in female compared with male patients (Fig 1, D).

evaluation, diagnosis, and treatment. In this study the soluble BAL cell proteome, primarily consisting of alveolar macrophages, was investigated with the aim to identify specific protein biomarkers of disease, as well as pathways of importance in early-stage disease pathology, with a focus on gender-specific disease phenotypes. The results showed significant differences in the soluble macrophage proteome between healthy smokers and smoking patients with early-stage COPD. Further investigation revealed that these differences were completely driven by the female population, with a subset of 19 proteins creating a highly significant model for classification of female nonsymptomatic smokers from female smokers with COPD ($P[\text{CV-ANOVA}] = 1.1 \times 10^{-7}$; Fig 1, B). While the classification model for women was highly robust, displaying a predictive power of 78%, the corresponding male classification model was nonsignificant ($P[\text{CV-ANOVA}] = 1.0$), with no altered proteins found in the univariate statistical analysis ($q < .05$). The observed

gender-specific differences were evident also when directly comparing smoking male and female patients with COPD (Fig 1, D). In contrast, no significant gender-based model could be created for the corresponding healthy smoker or healthy nonsmoker groups, indicating that the observed gender-related differences in the macrophage proteome are related to COPD disease phenotype and not to generic gender-specific differences. As such, the putative biomarkers identified in this study are only applicable to a female-dominated disease phenotype.

Cathepsin B (downregulated), ATP synthase (upregulated), and chaperonin (upregulated) were the most prominent disease marker candidates. All 3 proteins were highly significantly altered in univariate terms and represent the most important variables in the robust multivariate classification model. In addition, the 3 proteins were altered by 1.9- to 2.7-fold with low intragroup variance (Fig 1 and see Table E2), all important features for a potential formatting into a clinically applicable diagnostic assay.

TABLE II. Comparison of the most prominent variables (proteins) for separating the different groups of all comparisons relevant for this study

Spot ID	Protein (gene name)	Smoker vs COPD			COPD men vs women	Healthy vs exsmoker COPD
		C	♀	♂		
R^2/Q^2	All variables	0.73/0.19	0.94/0.70	0.86/0.17	0.96/0.30	0.82/0.05
$P(\text{CV-ANOVA})$		.05	8×10^{-5}	1.0	.05	.87
R^2/Q^2	Selected variables	0.50/0.45	0.81/0.78	0.49/0.42	0.76/0.71	0.60/0.50
$P(\text{CV-ANOVA})$		.0003	1×10^{-7}	.03	.0007	.0001
1561	<i>HSPD1 HNRNPK</i>	0.63	0.94		0.66	
1291	<i>ATP5H</i>	0.62	0.92		0.62	
447	<i>ATP6V1A</i>	0.63	0.90		0.71	
684	<i>ATP5B VIM</i>	0.64	0.90		0.71	
428	<i>GAA</i>	0.63	0.88		0.71	
490	<i>LTA4H</i>	0.61	0.79			0.66
1066	<i>PSMA6</i>		0.78		0.64	
1135	<i>SELENBP1 PRDX3</i>		0.77			
544	<i>HSPD1</i>		0.76			
1560	<i>VIM</i>		0.75			
1503	<i>UQCRC1</i>		0.74			
785	<i>AHCY SUCLA2</i>		0.73		0.65	
561	<i>CESI</i>		0.61		0.69	
1656	<i>CESI</i>				0.66	
766	<i>SLC25A24 TPP1</i>				0.67	
732	<i>SCP2</i>				0.70	
1182	<i>SPRYD4</i>				0.71	
392	<i>LMNA</i>					−0.61
780	<i>SLC25A24</i>				−0.66	
1394	<i>VIM</i>				−0.68	
397	<i>LMNA</i>					−0.61
1594	<i>VIM TUBB</i>				−0.76	
1397	<i>VIM</i>				−0.63	
1313	<i>LMNA</i>	−0.69		−0.67		
1566	<i>LMNA</i>	−0.66	−0.60	−0.66		
1486	<i>PDIA3</i>		−0.62			
1426	<i>CCT5</i>		−0.63		−0.60	
1094	<i>TPI1</i>		−0.65		−0.65	
1423	<i>GLB1 VIM</i>		−0.69			
1113	<i>CTSB</i>		−0.69			
1098	<i>CTSB</i>		−0.73		−0.65	

The absolute $p(\text{corr})$ values of 0.6 or greater from the respective multivariate analyses (OPLS) representing the variable importance are shown in the table. The qualitative model statistics are shown as R^2 , Q^2 , and P values from CV-ANOVA. The table is sorted by the strongest model (ie, comparison of female smokers vs female patients with COPD). Only spots with identified proteins are shown.

♀, Female healthy smokers versus female smokers with COPD (Fig 2, B); ♂, male healthy smokers versus male smokers with COPD (Fig 2, C); C, combined (healthy smokers vs smokers with COPD of both genders [Fig 2, A]).

The alterations of cathepsin B were confirmed by means of Western blotting, thereby providing proof of principle of the feasibility of transfer to lower-resolution separation methods suitable for clinical applications. Finally, the high correlation with the non-smoker classification model provides an additional independent validation step (Fig 3): all 3 putative biomarker proteins were prominent variables in driving the separation between healthy nonsmoking volunteers and exsmokers with COPD, further validating the relevance of these proteins in COPD pathology unrelated to smoking.

Female gender hormones have been suggested to have a protective role in the progression of COPD, with postmenopausal women at increased risk of disease.³ The demographics of the current study are in accordance with these observations: although all female patients with COPD had undergone menopause, only half of the nonsymptomatic smokers (9/17) were postmenopausal. In addition, none of the smokers with COPD

were taking hormone replacement therapy, whereas 33% of the nonsymptomatic smokers (3/9) did. However, the differentially expressed proteins were not dependent on menopausal status (see Fig E2).

The observed alterations in the BAL cell proteome primarily belong to 4 different cellular functions: energy metabolism (upregulated in patients with COPD), the lysosome (downregulated in patients with COPD), protein turnover (upregulated in patients with COPD), or structural proteins (downregulated in patients with COPD). Pathway analysis revealed that the lysosome activity pathway was significantly downregulated in women with COPD ($P = 3.1 \times 10^{-6}$), primarily driven by alterations of cathepsin B abundances. Downregulation of lysosomal function has been linked to autophagy, a conserved homeostatic pathway for protein degradation and regeneration. There is growing evidence in the literature that macroautophagy, one of the 3 main autophagic pathways, plays an important role in the pathogenesis

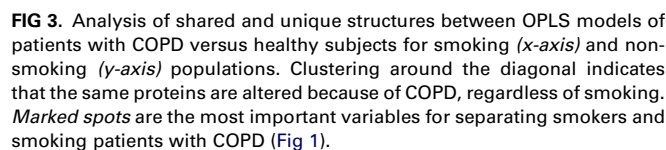
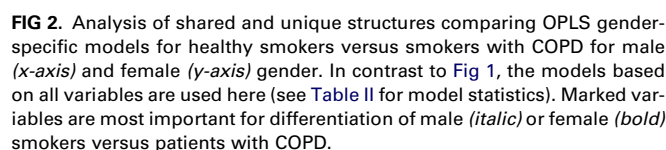


Figure 2 consists of two box plots, A and B, showing the relative band volume of α -actin and α -tubulin, respectively, in COPD and smoker groups. The y-axis for both plots is 'Relative band volume'.

Plot A: α -actin

- The y-axis ranges from 0.2 to 0.8.
- The x-axis categories are smoker1, COPD1, smoker2, and COPD2.
- Significance levels: $p = 0.01$ between smoker1 and COPD1; $p = 0.004$ between smoker2 and COPD2.

Plot B: α -tubulin

- The y-axis ranges from 0.05 to 0.15.
- The x-axis categories are smoker1, COPD1, smoker2, and COPD2.
- Significance levels: $p < 0.001$ between smoker1 and COPD1; $p < 0.00$ between smoker2 and COPD2.

FIG 4. Validation of 2D-DIGE results by means of Western blotting. **A**, Cathepsin B; **B**, lamin A/C.

differences in the pathophysiology of the distal lung, with less pronounced emphysema and thicker small airway walls in women compared with men matched in terms of lung function, age, and smoking history.⁵³

The COPD- and smoking-related inhibition of macroautophagy in alveolar macrophages reported by Monick et al.⁵² has also been associated with nonfunctional mitochondria and associated alterations in energy metabolism. As such, the significant upregulation of the oxidative phosphorylation pathway ($P = 5.4 \times 10^{-4}$) observed in these studies might represent a compensatory effect to the dysregulation of macroautophagy. The upregulation of enzymes involved in ATP production and consumption might also be an early indication of wasting syndrome often associated with more progressed COPD. Reports of nutritional depletion and loss of fat-free body mass,⁵⁴ as well as decreased ability to exercise and accumulation of lactic acid in muscle tissue, associated with COPD are indications of a deficiency in energy production.⁵⁵ Nutritional and muscle depletion in patients with COPD has been associated with female gender.^{56,57} Our observed upregulation of proteins involved in oxidative phosphorylation primarily in women with COPD thus supports the existence of a gender bias in COPD-associated dysregulation of energy metabolism. Upregulation of acid α -glucosidase could be a further indication of an increased need for additional and alternate routes of energy production.

In conclusion, this article makes an important contribution to the elucidation of putative molecular mechanisms underlying gender-specific differences in the pathophysiology of COPD. For the first time, the BAL cell proteome was systematically investigated for differences between male and female patients with COPD. The results showed clear differences in proteome alterations between men and women in early disease stages, with the female alveolar macrophages displaying significant phenotypic alterations due to the development of COPD, whereas no significant alterations were observed in men. Accordingly, this study gives clear indications that disease development in men and women is considerably different, indicating that future biomarker discovery efforts, as well as development of new diagnostic and treatment strategies in patients with COPD, would benefit from more of a gender-based focus.

We thank the research nurses Heléne Blomqvist, Margitha Dahl, and Gunnel de Forest, as well as biomedical analyst Benita Engvall for sample collection and preparation, as well as all of the volunteers for making the study possible.

Key messages

- Comparison of the BAL cell proteome of smokers and patients with COPD revealed significant alterations in proteins and pathways of female but not male patients with COPD, indicating gender-specific disease courses.
- Downregulation of the lysosomal pathway and upregulation of the oxidative phosphorylation pathway were observed and are possibly linked to dysregulation of macroautophagy and previously found phenotypic gender-specific differences.

REFERENCES

- Mathers CD, Boerma T, Ma Fat D. Global and regional causes of death. *Br Med Bull* 2009;92:7-32.
- Lopez AD, Murray CC. The global burden of disease, 1990-2020. *Nat Med* 1998;4:1241-3.
- Ohar J, Fromer L, Donohue JF. Reconsidering sex-based stereotypes of COPD. *Prim Care Respir J* 2011;20:370-8.
- Lundback B, Lindberg A, Lindstrom M, Ronmark E, Jonsson AC, Jonsson E, et al. Not 15 but 50% of smokers develop COPD?—report from the Obstructive Lung Disease in Northern Sweden studies. *Respir Med* 2003;97:115-22.
- Stockley RA, Mannino D, Barnes PJ. Burden and pathogenesis of chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009;6:524-6.
- Mannino DM, Homa DM, Akinbami LJ, Ford ES, Redd SC. Chronic obstructive pulmonary disease surveillance—United States, 1971-2000. *MMWR Surveill Summ* 2002;51:1-16.
- Jain NK, Thakkar MS, Jain N, Rohan KA, Sharma M. Chronic obstructive pulmonary disease: does gender really matter? *Lung India* 2011;28:258-62.
- Camp PG, Coxson HO, Levy RD, Pillai SG, Anderson W, Vestbo J, et al. Sex differences in emphysema and airway disease in smokers. *Chest* 2009;136:1480-8.
- Dransfield MT, Davis JJ, Gerald LB, Bailey WC. Racial and gender differences in susceptibility to tobacco smoke among patients with chronic obstructive pulmonary disease. *Respir Med* 2006;100:1110-6.
- Grydeland TB, Dirksen A, Coxson HO, Pillai SG, Sharma S, Eide GE, et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. *Eur Respir J* 2009;34:858-65.
- Townsend EA, Miller VM, Prakash YS. Sex Differences and Sex Steroids in Lung Health and Disease. *Endocr Rev* 2012;33:1-47.
- Gan WQ, Man SF, Postma DS, Camp P, Sin DD. Female smokers beyond the perimenopausal period are at increased risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Respir Res* 2006;7:52.
- Sorheim IC, Johannessen A, Gulsvik A, Bakke PS, Silverman EK, DeMeo DL. Gender differences in COPD: are women more susceptible to smoking effects than men? *Thorax* 2010;65:480-5.
- Ohlmeier S, Mazur W, Linja-Aho A, Louhelainen N, Ronty M, Toljamo T, et al. Sputum proteomics identifies elevated PI3R levels in smokers and mild-to-moderate COPD. *J Proteome Res* 2012;11:599-608.
- Casado B, Luisetti M, Iadarola P. Advances in proteomic techniques for biomarker discovery in COPD. *Expert Rev Clin Immunol* 2011;7:111-23.
- Chen H, Wang D, Bai C, Wang X. Proteomics-based biomarkers in chronic obstructive pulmonary disease. *J Proteome Res* 2010;9:2798-808.
- Waldburg N, Kahne T, Reisenauer A, Rocken C, Welte T, Buhling F. Clinical proteomics in lung diseases. *Pathol Res Pract* 2004;200:147-54.
- Verrills NM, Irwin JA, He XY, Wood LG, Powell H, Simpson JL, et al. Identification of novel diagnostic biomarkers for asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011;183:1633-43.
- Tewari AK, Popova-Butler A, El-Mahdy MA, Zweier JL. Identification of differentially expressed proteins in blood plasma of control and cigarette smoke-exposed mice by 2-D DIGE/MS. *Proteomics* 2011;11:2051-62.
- Nicholas BL, O'Connor CD, Djukanovic R. From proteomics to prescription—the search for COPD biomarkers. *COPD* 2009;6:298-303.
- Hu R, Ouyang Q, Dai A, Tan S, Xiao Z, Tang C. Heat shock protein 27 and cyclophilin A associate with the pathogenesis of COPD. *Respirology* 2011;16:983-93.
- Magi B, Bargagli E, Bini L, Rottoli P. Proteome analysis of bronchoalveolar lavage in lung diseases. *Proteomics* 2006;6:6354-69.
- York TP, van den Oord EJ, Langston TB, Edmiston JS, McKinney W, Webb BT, et al. High-resolution mass spectrometry proteomics for the identification of candidate plasma protein biomarkers for chronic obstructive pulmonary disease. *Biomarkers* 2010;15:367-77.
- Plymoth A, Lofdahl CG, Ekberg-Jansson A, Dahlback M, Broberg P, Foster M, et al. Protein expression patterns associated with progression of chronic obstructive pulmonary disease in bronchoalveolar lavage of smokers. *Clin Chem* 2007;53:636-44.
- Lee EJ, In KH, Kim JH, Lee SY, Shin C, Shim JJ, et al. Proteomic analysis in lung tissue of smokers and COPD patients. *Chest* 2009;135:344-52.
- de Torres JP, Casanova C, Pinto-Plata V, Varo N, Restituto P, Cordoba-Lanus E, et al. Gender differences in plasma biomarker levels in a cohort of COPD patients: a pilot study. *PLoS One* 2011;6:e16021.
- Larsson S, Nordenson A, Glader P, Yoshihara S, Linden A, Slinde F. A gender difference in circulating neutrophils in malnourished patients with COPD. *Int J Chron Obstruct Pulmon Dis* 2011;6:83-8.
- Barnes PJ. Alveolar macrophages as orchestrators of COPD. *COPD* 2004;1:59-70.
- Morjaria JB, Malerba M, Polosa R. Biologic and pharmacologic therapies in clinical development for the inflammatory response in COPD. *Drug Discov Today* 2010;15:396-405.
- Shaykhiyev R, Krause A, Salit J, Strulovici-Barel Y, Harvey BG, O'Connor TP, et al. Smoking-dependent reprogramming of alveolar macrophage polarization: implication for pathogenesis of chronic obstructive pulmonary disease. *J Immunol* 2009;183:2867-83.
- Lofdahl JM, Wahlstrom J, Skold CM. Different inflammatory cell pattern and macrophage phenotype in chronic obstructive pulmonary disease patients, smokers and non-smokers. *Clin Exp Immunol* 2006;145:428-37.
- Stocks J, Quanjer PH. Reference values for residual volume, functional residual capacity and total lung capacity. *ATS Workshop on Lung Volume Measurements. Official Statement of The European Respiratory Society. Eur Respir J* 1995;8:492-506.
- Eklund A, Blaschke E. Relationship between changed alveolar-capillary permeability and angiotensin converting enzyme-activity in serum in sarcoidosis. *Thorax* 1986;41:629-34.
- Lofdahl JM, Cederlund K, Nathell L, Eklund A, Skold CM. Bronchoalveolar lavage in COPD: fluid recovery correlates with the degree of emphysema. *Eur Respir J* 2005;25:275-81.
- Storey J. A direct approach to false discovery rates. *J R Stat Soc B* 2002;64:479-98.
- Jackson JE. A user's guide to principal components. New York: John Wiley; 1991.
- Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). *J Chemometrics* 2002;16:119-28.
- Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH, et al. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. *Electrophoresis* 2000;21:3666-72.
- Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J Am Soc Mass Spectrom* 1994;5:976-89.
- Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4:P3.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44-57.
- Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;37:1-13.
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 1999;27:29-34.
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012;40(Database issue):D109-14.
- Rennard SI, Ghafouri M, Thompson AB, Linder J, Vaughan W, Jones K, et al. Fractional processing of sequential bronchoalveolar lavage to separate bronchial and alveolar samples. *Am Rev Respir Dis* 1990;141:208-17.
- Schildge J, Nagel C, Grun C. Bronchoalveolar lavage in interstitial lung diseases: does the recovery rate affect the results? *Respiration* 2007;74:553-7.
- Silva E, O'Gorman M, Becker S, Auer G, Eklund A, Grunewald J, et al. In the eye of the beholder: does the master see the same spots as the novice? *J Proteome Res* 2010;9:1522-32.
- Trygg J, Wold S. Orthogonal Projections to Latent Structures (OPLS). *J Chemometrics* 2002;16:119-28.
- Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, et al. Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinformatics* 2009;10:136.
- Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res* 2011;39(Database issue):D214-9.

51. Haspel JA, Choi AM. Autophagy: a core cellular process with emerging links to pulmonary disease. *Am J Respir Crit Care Med* 2011;184:1237-46.
52. Monick MM, Powers LS, Walters K, Lohan N, Zhang M, Gerke A, et al. Identification of an autophagy defect in smokers' alveolar macrophages. *J Immunol* 2010; 185:5425-35.
53. Martinez FJ, Curtis JL, Sciurba F, Mumford J, Giardino ND, Weinmann G, et al. Sex differences in severe pulmonary emphysema. *Am J Respir Crit Care Med* 2007;176:243-52.
54. Decramer M, De Benedetto F, Del Ponte A, Marinari S. Systemic effects of COPD. *Respir Med* 2005;99(suppl B):S3-10.
55. Green HJ, Bombardier E, Burnett M, Iqbal S, D'Arsigny CL, O'Donnell DE, et al. Organization of metabolic pathways in vastus lateralis of patients with chronic obstructive pulmonary disease. *Am J Physiol Regul Integr Comp Physiol* 2008;295: R935-41.
56. Verhage TL, Heijdra Y, Molema J, Vercoulen J, Dekhuijzen R. Associations of muscle depletion with health status. Another gender difference in COPD? *Clin Nutr* 2011;30:332-8.
57. Vermeeren MA, Creutzberg EC, Schols AM, Postma DS, Pieters WR, Roldaan AC, et al. Prevalence of nutritional depletion in a large out-patient population of patients with COPD. *Respir Med* 2006;100:1349-55.

METHODS

Chemicals

All reagents for buffers were purchased from GE Healthcare (Uppsala, Sweden) or Sigma-Aldrich (Munich, Germany) and were of analytic or electrophoretic grade. Buffers and solutions were prepared by using MilliQ water (18 M Ω ; Millipore, Solna, Sweden).

Subjects and study design

Chest radiography and clinical examination did not reveal any other significant disease than COPD. Clinical data concerning symptoms, medications (including oral contraceptives or estrogen replacement), smoking history, and additional exposures relevant for COPD were recorded by using a questionnaire. Group statistics for the subset of subjects included in the final statistical analyses of the current study are presented in Table I, and differences were calculated by using the Student *t* test (Microsoft Excel 2007, 2-tailed, homoscedastic). The study was approved by the Stockholm Regional Ethical Board (case no. 2006/959-31/1), and all volunteers provided informed written consent.

Sample collection

Samples were collected by means of BAL, as previously described,^{E1,E2} from the middle lobe bronchus by using 5 $\times$ 50 mL of PBS (37°C). The combined aspirates were filtered and centrifuged at 400g for 5 minutes at 4°C. Cells were resuspended in 7.5 mL of PBS and divided for differential cell counts and proteomics. For proteomics analysis, aliquots of 3 million cells were placed in pellets by means of centrifugation at 13,000g for 2 minutes, resuspended in 1 mL of freeze medium (RPMI 1640, 10% dimethyl sulfoxide, and 1% protease inhibitor [Sigma]), and frozen at -80°C until further preparation and analysis. Differential cell counts were performed with a Bürker chamber and trypan blue staining.

2D-DIGE

BAL fluid cells in freeze medium were placed in pellets by means of centrifugation for 2 minutes at 13,000g, washed twice with PBS containing protease inhibitor (Complete; Roche, Bromma, Sweden), and resuspended in 150 to 250 μL of lysis buffer (8 mol/L urea, 2 mol/L thiourea, 4% Chaps, and 33 mmol/L Tris). Cells were ground 2 $\times$ 1 minute at 20 Hz by using a Qia-gen tissue lyzer II (Sollentuna, Sweden) and subsequently incubated at room temperature for 1 hour for complete cell lysis. After centrifugation for 10 minutes at 13,000g, the protein concentration in the supernatant was determined by using the Bradford method (Bio-Rad, Sundbyberg, Sweden), and samples were diluted to 1 mg/mL and frozen at -80°C until further analysis.

Fifty micrograms of each sample was labeled with 400 pmol of DIGE minimal labeling dye according to the supplier's recommendations. A triplexed experimental design applying coseparation of a Cy2-labeled pooled internal standard and 2 samples randomly assigned to Cy3 or Cy5 labeling was used. The internal standard was created from pooling of most of the samples (83%) to be analyzed, consisting of equal amounts of protein from each group. For lysis, labeling, and gel runs, samples were repetitively randomized to avoid systematic bias caused by analysis batches.

After CyDye labeling, 2 samples and 1 internal standard were pooled, the sample volume was adjusted to 350 μL with lysis buffer 2 (8 mol/L urea, 2 mol/L thiourea, 4% Chaps, and 4 μL of a 1% bromophenol blue solution [wt/vol]), and 3.5 μL (1%) of IPG buffer (pH 4-7; GE Healthcare), was added.

Isoelectric focusing was performed with 18-cm strips with a pH range of 4 to 7 (GE Healthcare). Samples were applied to the strip by means of rehydration loading overnight, and isoelectric focusing was performed overnight (IPGphor II) at voltage gradients of 200 V for 1 hour, 200 to 3500 V for 3 hours, and 3500 V for 19 hours, resulting in a run of 68 kVhrs.

After isoelectric focusing, strips were equilibrated 2 times for 15 minutes in 6 mol/L urea, 75 m mol/L Tris-HCl (pH 8.8), 2% SDS, 30% glycerol, and 0.002% bromophenol blue. For reduction and alkylation, 10 mg/mL dithiothreitol was added in the first and 25 mg/mL iodoacetamide in the second equilibration steps. SDS-PAGE was performed on lab-casted 10% Tris-glycine gels by using Bio-Rad (Munich, Germany) acrylamide solution

(10% T and 3.3% C) and recipe (375 mmol/L Tris [pH 8.8], 0.1% SDS, 0.05% N,N,N',N'-tetramethylethylenediamine [vol/vol], and 0.05% ammonium persulfate [wt/vol]). Gels were run in an Ettan Dalt 6 chamber (GE Healthcare) with running buffer containing 25 mmol/L Tris, 192 mmol/L glycine, and 0.1% SDS (1 mA per gel; 1 W per gel for 45 minutes, 60 mA per gel; 15 W per gel for approximately 5 hours).

Image acquisition was performed with an FLA Typhoon 9000 laser scanner (Cy2: laser 473 nm, emission filter 530BP20; Cy3: laser 532 nm, emission filter 570BP20; Cy5: laser 635 nm, emission filter 665LP). Photomultiplier tube voltages were adjusted to maximize the dynamic range of pixel intensities in the gel images while avoiding saturation of the highest-intensity protein spots.

Quantitation and statistical analyses

In addition to *P* values and FDR-corrected *P* values, a fold change of at least 1.3 was used as an additional stringency criterion for variable selection.

Protein identification by means of mass spectrometry

For preparative gels for in-gel trypsin digestion and protein identification by using liquid chromatography–tandem mass spectrometry, 150 to 250 μg of protein was loaded on one IPG strip and separated by using 2-DE, as described above. After SDS-PAGE, gels were stained with silver according to the method of Yan et al.^{E3}

Differentially expressed protein spots of interest were excised, and the gel plugs were destained with 30 mmol/L potassium ferricyanide/100 mmol/L sodium thiosulfate 1:1. After in-gel trypsin digestion overnight (200 ng trypsin per gel spot), peptides were extracted 3 times with 1% trifluoroacetic acid/50% acetonitrile, and the combined extracts were dried in a Speed-Vac.^{E4} After reconstitution in 30 μL of 2% acetic acid, peptides were analyzed by means of nano-LC-Orbitrap-MS. Nanoflow liquid chromatography was performed on a Waters nano-Acquity UPLC (Eschborn, Germany) by using a Symmetry C18 trapping column (5 μm ; 180 μm $\times$ 20 mm; flow, 5 $\mu\text{L}/\text{min}$) and a BEH130C18 peptide column (1.7 μm , 100 μm $\times$ 100 mm) at a flow rate of 750 nL/min. A gradient program with 0.1% formic acid as solvent A and acetonitrile containing 0.1% formic acid as solvent B was used as follows: 3 minutes at 97% A, 3 to 5 minutes at 80% A, 5 to 45 minutes at 40% A, 45 to 48 minutes at 20% A, 48 to 50 minutes at 3% A, 50.01 minutes at 97% A, and 15 minutes of equilibration. The LC was coupled to the mass spectrometer by using a nano-spray ion source operated in positive ionization mode with an ionization voltage of 1.4 kV. The Orbitrap mass spectrometer (Thermo, Bremen, Germany) recorded full-scan spectra with a resolution of 30,000 (FWHM), and the collision energy for MS/MS experiments was set to 30% (arbitrary unit, Xcalibur 2.0 SR2, Thermo). The 3 most intensive signals, excluding singly charged species, were fragmented by using helium as a collision gas. The damping gas used was nitrogen supplied by a CMC nitrogen generator (CMC Instruments, Eschborn, Germany). The Orbitrap was calibrated with the manufacturer's calibration solution (consisting of caffeine, the tetrapeptide MRFA, and ultramark), allowing for accurate mass measurement with errors of less than 5 ppm. For identification of proteins by using database searching, Proteome Discoverer 1.0 was used with Sequest^{E5} and the human FASTA database (2006). Search parameters were 1 missing cleavage for trypsin, iodoacetamide derivatization of cysteine residues, methionine oxidation as dynamic modification, an FDR of 0.01, a mass error of 10 ppm for mass spectra, and 0.8 Da for the fragment ions.

Validation of proteins by means of Western blotting

Further validation of the 2D-DIGE results by using Western blotting was performed for 2 proteins differentially expressed between smokers and patients with COPD: lamin A/C and cathepsin B. The samples used for validation (7 smokers and 7 patients with COPD) were the nearest to the group mean values of the 2D-DIGE gel spot volumes (lamin A/C, spot 1566; cathepsin B, spot 1113) to ensure a representative selection. An amount of 10 μg of protein from each sample lysate was mixed with lithium dodecyl sulfate sample buffer and separated on NUPAGE Bis-Tris mini gels (10%, 15 wells, 0.1 cm; Life

Technologies, Stockholm, Sweden) with 3-(N-morpholino)propanesulfonic acid running buffer and run at a constant voltage of 125 V for approximately 1.5 hours. Proteins were wet blotted (Mini Trans-Blot, Bio-Rad) onto nitrocellulose membranes (GE Healthcare) by using a buffer containing 25 mmol/L Tris, 192 mmol/L glycine, and 20% EtOH for 1 hour at 100 V and 350 mA. Membranes were blocked in 2% blocking agent (GE Healthcare) in PBST (PBS with 0.1 Tween 20) incubated with primary antibodies (rabbit anti-actin antibody [Sigma A2066, dilution 1:200], mouse anti-cathepsin B [Sigma, dilution 1:400], and mouse anti-lamin A/C [BD Biosciences, 1:500]) for 1 hour at room temperature in blocking solution and then incubated with secondary ECL Plex CyDye-labeled antibodies (GE Healthcare, Cy5-labeled anti-mouse antibody; Cy3-labeled anti-rabbit antibody, diluted 1:2500) for 1 hour in PBST. Blot image acquisition was performed by using a Typhoon 9000 fluorescence scanner. Image Quant 7.0 (GE Healthcare) was used to calculate band volumes by using rolling-ball background subtraction. The corresponding band densities of actin were used for normalization of protein abundances between samples.

REFERENCES

- E1. Eklund A, Blaschke E. Relationship between changed alveolar-capillary permeability and angiotensin converting enzyme-activity in serum in sarcoidosis. *Thorax* 1986;41:629-34.
- E2. Lofdahl JM, Cederlund K, Nathell L, Eklund A, Skold CM. Bronchoalveolar lavage in COPD: fluid recovery correlates with the degree of emphysema. *Eur Respir J* 2005;25:275-81.
- E3. Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH, et al. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. *Electrophoresis* 2000;21:3666-72.
- E4. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* 1999;20:601-5.
- E5. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J Am Soc Mass Spectrom* 1994;5:976-89.

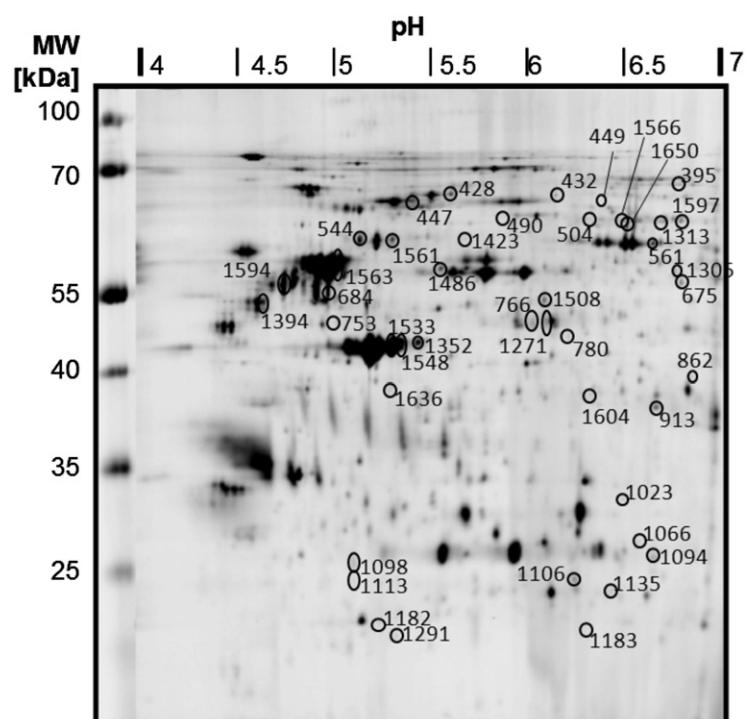


FIG E1. Representative gel image of the 2DE-separated soluble proteins from BAL cells of a pool of healthy volunteers, smokers, and patients with COPD (internal standard). The spots indicated are further described in Fig 2 and Tables III and E2 and represent the most important findings in this study.

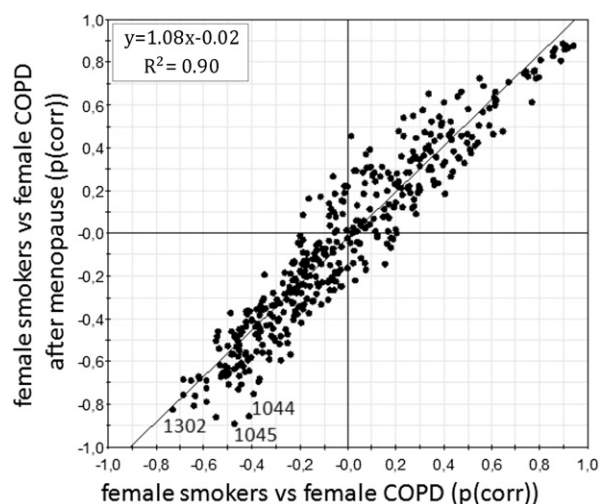


FIG E2. SUS plot comparing the predictive component of the models comparing all female smokers with female patients with COPD (*x-axis*; $R^2 = 0.94$, $Q^2 = 0.70$) and female postmenopausal smokers versus female postmenopausal patients with COPD (*y-axis*; $R^2 = 0.97$, $Q^2 = 0.72$). As evident from the diagonal distribution of the variables, menopause does not appear to cause any substantial alterations to the variable contribution. The protein spots that are most affected by the removal of premenopausal women from the model (furthest away from the diagonal) include prohibitin (spot 1044), cathepsin D (spot 1044, 1045), and hexosaminase B (spot 1302). All 3 proteins become more prominent driving factors when nonmenopausal women are excluded, as indicated by the higher p(corr) values on the y-axis.

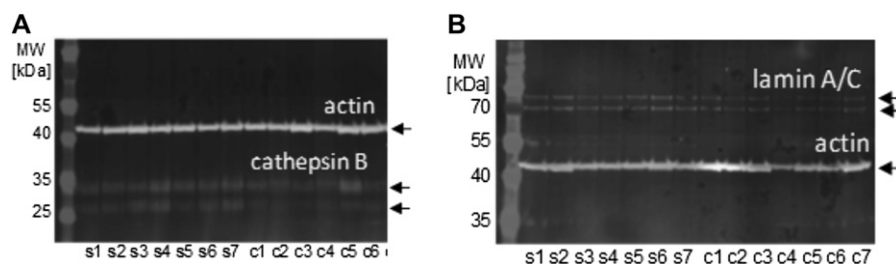


FIG E3. Western blot images for validation of differentially expressed proteins (**A**, cathepsin B; **B**, lamin A/C). Experimental molecular weights (*MW*) obtained by using 2D-DIGE for cathepsin were 32 kDa for spot 1026 and 25 kDa for spots 1098 and 1113. Molecular weights obtained for lamin A/C include 68 kDa for spots 392, 395, and 397 and 64 kDa for spots 1313, 1587, 1600, and 1619. The dual molecular weights observed in the Western blots for each of the proteins (*arrowheads*) correspond with those measured by using 2D-DIGE for both proteins. A single band was detected for actin (45 kDa), which was used for normalization of the band volumes.

TABLE E1. Significant differences of clinical parameters between study groups

Parameter	Smoker vs COPD, combined	Smoker vs COPD, female	Smoker vs COPD, male	COPD gender	Healthy vs exsmoker COPD
Age	NS	NS	NS	NS	NS
BMI	NS	NS	NS	NS	NS
Recovered BAL (%)	0.003	0.002	NS	NS	0.006
Macrophages (%)	NS	NS	NS	NS	NS
Neutrophils (%)	NS	NS	NS	NS	0.0002
Pack years	NS	NS	NS	NS	NS
FEV ₁ (%)	0.04	0.0002	1.6×10^{-7}	0.02	0.0003
FEV ₁ /VC ratio (%)	4×10^{-8}	0.0001	0.0002	NS	4×10^{-7}
FEV ₁ /FVC ratio (%)	3×10^{-14}	3×10^{-8}	3×10^{-7}	NS	5×10^{-6}

BMI, Body mass index; NS, not significant.

TABLE E2. Identified proteins of interest and their main functions

Spot ID	Protein name and function	Alterations			
		C	♀	♂	G
395	Lamin A/C: part of the nuclear lamina	−1.2x			
428	Acid α-glucosidase: degradation of glycogen in lysosome	+1.3x	+1.6x		f1.4x
432	Vesicle transport-related protein isof b: transport ER ↔ Golgi	−1.1x		−1.2x	
447	ATPase: energy metabolism	+1.3x	+1.6x		f1.4x
490	Leukotriene A ₄ hydrolase: biosynthesis of leukotriene B ₄ /aminopeptidase activity	+1.3x			
544	Chaperonin: protein folding in mitochondria	+1.2x			
561	Carboxylesterase 1 isoform c: detoxification of xenobiotics/activation of ester and amide prodrugs (ER)				f2.0x
675	Septin 11: filament forming GTPase/cytokinesis	−1.2x	−1.2x		
684	ATP synthase: energy metabolism	+1.5x	+1.9x		f1.4x
766	(A) Solute carrier fam 25 mem 24 isof 1: mitochondrial carrier (B) Tripeptidyl-peptidase I: lysosomal proteinase product degradation				f1.3x
780	Solute carrier fam 25 mem 24 isof 1: mitochondrial carrier				m1.3x
1066	Proteasome α-6 subunit: cleavage of peptides		+1.7x		
1094	Trios phosphate isomerase: glyconeogenesis/glycolysis		−1.4x		m1.4x
1098	Cathepsin B: lysosomal protein degradation	−1.7x	−2.7x		
1113	Cathepsin B: lysosomal protein degradation		−2.5x		
1135	(A) Se binding protein: intra-Golgi protein transport (B) Peroxiredoxin 3: redox regulation; activation of NF-κB	+1.5x	+1.6x		
1182	SPRY domain-containing protein 4: unknown				f1.4x
1271	(A) Tripeptidyl peptidase I: lysosomal proteinase product degradation (B) Solute carrier fam 25, member 24: mitochondrial carrier			−1.2x	
1291	ATP synthase: energy metabolism	+1.7x	+2.6x		
1305	Leucine aminopeptidase 3: processing/turnover of intracellular proteins				m1.6x
1313	Lamin A/C: part of the nuclear lamina	−1.3x		−1.3x	
1394	Vimentin: class-III intermediate filaments				m1.3x
1423	(A) Galactosidase β: formation of elastin fibers (B) Vimentin: class-III intermediate filaments	−1.2x	−1.5x		
1486	Disulfide isomerase A3: rearrangement of disulfide bonds	−1.3x			
1508	(A) sterol carrier protein 2: transfer of lipids between membranes (B) Mitochondrial processing peptidase β subunit: cleaves presequences from mitochondrial protein precursors	−1.2x			
1561	Chaperonin, HSP60: protein folding in mitochondria	+1.4x	+1.9x		
1566	Lamin A/C: part of the nuclear lamina	−1.4x	−1.4x	−1.4x	
1594	(A) Vimentin: class-III intermediate filaments (B) Tubulin beta: major constituent of microtubules				m1.3x
1636	Alpha-glycosidase II, isof 2: synthesis of glycoproteins			−1.2x	

The right column indicates for which of the OPLS models presented in Fig 2 the respective protein was the most important variable in driving group separation.

♀, Female healthy smokers versus female smokers with COPD (Fig 2, B); ♂, male healthy smokers versus male smokers with COPD (Fig 2, C); C, combined (healthy smokers vs smokers with COPD of both genders [Fig 2, A]); *fx/mx*, fold changes from univariate analysis, higher abundance in female subjects (*f*) or male subjects (*m*); G, gender model (male patients with COPD vs female patients with COPD [Fig 2, D]); *−x/+x*, fold changes from univariate analysis, upregulated (+) or downregulated (−) in patients with COPD.

TABLE E3. Differentially expressed proteins discussed in the main document (see Figs 2 and E1 and Tables II and III) with score, sequence coverage, and number of identified peptides from liquid chromatography–mass spectrometry and database search, as well as experimental and theoretic isoelectric points and molecular weights

Spot ID	Protein name	Gene name	Accession no.	Score	Sequence coverage (%)	No. of peptides	Experimental (E)/theoretic (T); MW/pl*
392	Lamin A, C isoform 2	<i>LMNA</i>	P02545	85	34	19	E: 68/6.8 T: 72.2/6.4
395	Lamin A, C isoform 1	<i>LMNA</i>	P02545	137	36	25	E: 68/6.8 T: 72.2/6.4
397	Lamin A, C isoform 1	<i>LMNA</i>	P02545	194	34	20	E: 68/6.9 T: 72.2/6.4
428	Acid α -glucosidase	<i>GAA</i>	P10253	46	3	3	E: 67/5.7 T: 105/5.6
432	Vesicle transport-related protein isoform b	<i>SCFD1</i>	Q8WVM8	33	14	6	E: 66/6.2 T: 72.2/5.9
447	ATPase H ⁺ transporting lysosomal 70 kDa V1 subunit A, isoform 1	<i>ATP6V1A</i>	P38606	113	29	16	E: 66/5.5 T: 68.3/5.3
490	Leukotriene A ₄ hydrolase	<i>LTA4H</i>	P09960	194	28	14	E: 64/6.0 T: 69.1/5.8
544	Chaperonin	<i>HSPD1</i>	P10809	190	28	16	E: 61/5.2 T: 85/5.2
561	Carboxylesterase 1, isoform c	<i>CES1</i>	P23141	147	26	13	E: 61/6.7 T: 60.6/6.1
675	Septin 11	<i>SEPT11</i>	Q9NVA2	113	22	10	E: 56/4.8 T: 49.3/6.4
684	ATP synthase H ⁺ transporting mitochondrial F1 complex, β subunit	<i>ATP5B</i>	P06576	395	35	15	E: 56/5.1 T: 53.5/5.1
	Vimentin	<i>VIM</i>	P08670	288	41	22	T: 53.5/5.1
732	Sterol carrier protein 2 isoform 1	<i>SCP2</i>	P22307	45	10	4	E: 52/6.3 T: 59.0/6.4
766	Solute carrier family 25, member 24, isoform 1	<i>SLC25A24</i>	Q6NUK1	74	20	8	E: 49/6.1 T: 53.3/6.0
	tripeptidyl-peptidase I	<i>TPPI</i>	O14773	65	8	3	T: 39.8/5.7
780	Vesicle amine transport protein 1	<i>TPPI</i>	O14773	65	8	3	T: 39.8/5.7
780	Solute carrier family 25, member 24, isoform 1	<i>SLC25A24</i>	Q6NUK1	171	37	16	E: 47/6.3 T: 53.3/6.0
785	S-adenosylhomocysteine hydrolase	<i>AHCY</i>	P23526	55	28	12	E: 46/6.2
	Succinate-CoA ligase ADP-forming β subunit	<i>SUCLA2</i>	Q9P2R7	49	25	12	T: 27.6/5.9 T: 44.6/5.6
1066	Proteasome α 6 subunit	<i>PSMA6</i>	P60900	58	41	7	E: 27/6.6 T: 35.4/6.3
1094	Triose phosphate isomerase 1	<i>TPI1</i>	P60174	60	32	7	E: 27/6.7 T: 26.5/6.5
1098	Cathepsin B	<i>CTSB</i>	P07858	40	19	3	E: 26/5.2 T: 27.8/5.2
1113	Cathepsin B	<i>CTSB</i>	P07858	100	20	12	E: <25/5.2 T: 27.8/5.2
1135	Selenium binding protein 1	<i>SELENBP1</i>	Q13228	59	17	6	E: <25/6.5
	Peroxiredoxin 3 isoform b	<i>PRDX3</i>	P30048	42	26	5	T: 21.5/5.8
1182	SPRY domain-containing protein 4	<i>SPRYD4</i>	Q8WW59	52	31	7	E: <25/5.4 T: 23.1/6.4
1271	Tripeptidyl-peptidase I	<i>TPPI</i>	O14773	59	8	3	E: 50/6.2
	Solute carrier family 25, member 24, isoform 1	<i>SLC25A24</i>	Q6NUK1	40	11	4	T: 53.4/6.0 T: 53.3/6.0
1291	ATP synthase H ⁺ transporting mitochondrial F0 complex, subunit d, isoform b	<i>ATP5H</i>	O75947	22	28	4	E: <25/5.4 T: 18.3/5.4
1305	Leucine aminopeptidase 3	<i>LAP3</i>	P28838	204	37	16	E: 58/6.8 T: 56.2/8.0
1313	Lamin A, C isoform 2	<i>LMNA</i>	P02545	233	48	30	E: 63/6.7 T: 72.2/6.4

(Continued)

TABLE E3. (Continued)

Spot ID	Protein name	Gene name	Accession no.	Score	Sequence coverage (%)	No. of peptides	Experimental (E)/theoretic (T); MW/pI*
1394	Vimentin	<i>VIM</i>	P08670	395	41	23	E: 55/4.8 T: 53.5/5.1
1397	Vimentin	<i>VIM</i>	P08670	270	36	18	E: 57/4.9 T: 53.5/5.1
1423	Galactosidase β 1	<i>GLB1</i>	P16278	73	20	12	E: 61/5.8
	Vimentin	<i>VIM</i>	P08670	35	28	11	T: 60.0/5.6 T: 53.5/5.1
1426	Chaperonin-containing TCP1 subunit 5 (ϵ)	<i>CCT5</i>	P48643	110	34	22	E: 61/5.7 T: 59.5/5.4
1486	Protein disulfide isomerase-associated 3	<i>PDIA3</i>	P30101	11	5	2	E: 59/5.9 T: 54.3/5.6
1503	Ubiquinol-cytochrome c reductase core protein I	<i>UQCRC1</i>	P31930	98	19	10	E: 52/5.5 T: 49.1/5.4
1508	Sterol carrier protein 2 isoform 1 proprotein	<i>SCP2</i>	P22307	122	15	9	E: 54/6.2
	Mitochondrial processing peptidase β subunit	<i>PMPCB</i>	O75439	42	18	8	T: 50.0/5.9 59/6.4
1560	Vimentin	<i>VIM</i>	P08670	7	5	2	E: 57/5.2 T: 53.5/5.1
1561	Chaperonin	<i>HSPD1</i>	P10809	204	28	14	E: 61/5.4 T: 58.0/5.2
1566	Lamin A, C isoform 2	<i>LMNA</i>	P02545	256	33	20	E: 64/6.5 T: 72.2/6.4
1594	Vimentin	<i>VIM</i>	P08670	387	52	30	E: 56/4.9
	Tubulin β polypeptide	<i>TUBB</i>	P07437	133	38	14	T: 53.5/5.1 49.7/4.8
1656	Carboxylesterase 1, isoform c	<i>CES1</i>	P23141	137	19	9	E: 61/6.7 T: 60.1/6.1

MW, Molecular weight; pI, isoelectric point.

*Experimental values are approximations from the 2DE gels.