Biomarkers, 2011; 16(7): 567–576

ISSN 1354-750X print/ISSN 1366-5804 online
DOI: 10.3109/1354750X.2011.608855

informa
healthcare

RESEARCH ARTICLE

Proteomic profiling of human respiratory epithelia by iTRAQ reveals biomarkers of exposure and harm by tobacco smoke components

Keith Sexton[1], Dominique Balharry[1], Paul Brennan[2], James McLaren[1], Ian A. Brewis[2], and Kelly A. BéruBé[1]

[1]Cardiff School of Biosciences, Cardiff University, Museum Avenue, Cardiff Wales, CF10, UK and [2]Department of Infection, Immunity and Biochemistry, School of Medicine, Cardiff University, Heath Park, Cardiff Wales, CF14 4XN, UK

Abstract
Historically, it has been challenging to go beyond epidemiology to investigate the pathogenic changes caused by tobacco smoking. The EpiAirway-100 (MatTek Corp., Ashland, MA) was employed to investigate the effects of cigarette smoke components. Exposure at the air-liquid-interface represented particle and vapour phase components of cigarette smoke. A proteomic study utilising iTRAQ labelling compared expression profiles. The correlative histopathology revealed focal regions of hyperplasia, hypertrophy, cytolysis and necrosis. We identified 466 proteins, 250 with a parameter of two or more peptides. Four of these proteins are potential markers of lung injury and three are related to mechanistic pathways of disease.

Keywords: Biomarkers, lung, tobacco smoke, isobaric labeling, tandem mass spectrometry

Introduction

Cigarette smoking is a major public health concern worldwide and is responsible for 5 million deaths each year (Lancet, 2009). It is associated with a number of pulmonary diseases, most notably chronic obstructive pulmonary disease (COPD) and lung cancer, but is also linked to cardiac and cerebrovascular diseases (Jemal et al., 2010). Tobacco smoke contains numerous toxic and carcinogenic compounds that interact with DNA, lipids and proteins in multiple organs to alter their normal physiological activities, and ultimately lead to adverse health effects (Hoffmann et al., 2001). Therefore, it is of primary importance to be able to characterise the insult of cigarette smoking through molecular biomarkers that can lead to better understanding of an individual's susceptibility to harmful health effects.

Lung tissue samples represent the best choice for understanding direct target organ effects. Alterations of structure and/or expression of proteins in lung cells/tissues can be very useful in providing diagnoses to clinical conditions. With the advent of human tissue-engineering, it is now possible to generate high-fidelity tissue-constructs, based on the targeted interactions of organ-specific cells, cultured in bioreactors that favour growth and differentiation in the air-liquid-interface (ALI); emulating the natural environment of the native lung *in vivo* (Berube et al., 2010). The commercial respiratory epithelial model EpiAirway™ (MatTek Corp., Ashland, MA), represents a fully-differentiated, 3 dimensional (3D), muco-ciliary phenotype that mimics the *in situ*, proximal, airway tissue found in humans (Bérubé et al., 2006; Balharry et al., 2008; Sexton et al., 2008; BéruBé et al., 2009).

The objective of this investigation was to assess the effect of tobacco smoke components (TSC) on the global expression of proteins in human respiratory epithelia (EpiAirway™ lung constructs) for biomarker discovery that may be of clinical value regarding tobacco-related diseases. A select range of TSC were utilised as surrogates of tobacco smoke, and represented drivers of the

Address for Correspondence: Keith Sexton, Cardiff School of Biosciences, Cardiff University, Museum Avenue, Cardiff Wales, UK. CF10 3US. E-mail: Sextonk@cardiff.ac.uk

(Received 16 May 2011; revised 25 July 2011; accepted 26 July 2011)

Abbreviations

BCH, Basal cell hyperplasia
Ca, Calcium
CI, Confidence interval
COPD, Chronic obstructive pulmonary disease
ELM, EpiAirway lung model
FC, Fold change
GPS, Global proteome server
HRP, Horseradish peroxidase
Mg, Magnesium
MMTS, Methyl methane-thiosulfonate
Nano-LC, Nano-liquid chromatography
NHBE, Normal human bronchial epithelium
NHT/B, Normal human tracheal/bronchial
PVDF, Polyvinylidenedifluoride
RH, Regional hypertrophy
ROI, Reactive oxygen intermediates
SCX, Strong cation exchange
TB, Toluidine Blue
TCEP, Tris(2-carboxyethyl)phosphine
TD_{20}, Toxic dose 20% (dose required to kill 20% of cells)
TS, Tobacco smoke
TSC, Tobacco smoke components

differential pathways of lung toxicity: nicotine [thrombogenic], cadmium [cytotoxic], urethane and formaldehyde [xenobiotic metabolism/mutagenic] (Balharry et al., 2008; Sexton et al., 2008). Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) and nano-liquid chromatography (nano-LC) proteomic methodology and matrix-assisted laser desorption/ionization tandem mass spectrometry (MALDITOF/TOF) was utilised to identify and quantify differentially expressed lung cell proteins between TSC-exposed and non-exposed EpiAirway™ lung constructs; select proteins were further confirmed by immunoblot analysis. The correlative histopathology of the EpiAirway™ cells, pre- and post-exposure to TSC, provided a morphological assessment of toxicity. A combination of conventional and 'omic'-technologies have enabled a better resolution of the pathogenesis of lung injury following exposure to tobacco smoke and the identification of associations between molecular changes and clinical end-points.

Methods

Cells, culture and exposure conditions

The EpiAirway cell cultures were transported from MatTek in the USA in a 24 well plate format. The cells were equilibrated at 37°C with 4.5% CO_2 for 24 h. Following this culture preconditioning, acute exposure (24 h) of the EpiAirway lung model (ELM) to TSC solutions was carried out at the air-liquid interface. TSC solutions were made up in sterile PBS (+ Mg and Ca) and warmed to 37°C before dosing. Each insert (surface area 1 cm^2) was apically dosed with 100 µl warm TSC solution. A concentration range for each TSC was determined from previous studies described in (Balharry et al., 2008). The concentrations (TD_{20}; the dose level at which a substance produces a toxic effect of 20%) required for the proteomic study were as follows: Nicotine: 30 mM; Cadmium (cadmium chloride): 0.1 mM; Urethane: 70 mM; Formaldehyde: 7 mM. Each dose was replicated ($n = 5$). After 24 h the toxin (TSC) was removed from the surface and the epithelial layer was stored as a frozen pellet prior to processing for proteomics or fixed in glutaraldehyde for semi-thin sectioning (see below).

Fixation of ELM for semi-thin sectioning

The ELM were excised from the Millicell plastic insert and submerged in cold (4°C) glutaraldehyde (3%) fixative for 15 min. The ELM was then removed and immersed in more fresh fixative for 1 h at 4°C. The gluteraldehyde fixative was replaced with phosphate buffer and washed overnight (~12 h), at 4°C.

Processing of ELM

Post-fixation was carried out by osmicating (1% osmium tetroxide in phosphate buffer) for 60 min at 4°C. The epithelial layer was then passed through a series of graded alcohols. Once dehydrated, the lung epithelial layer samples were immersed in propylene oxide for 30 min (Glauert and Lewis, 1998). This was followed by overnight rotation in a 50/50 mix of propylene oxide and Araldite CY212. The epithelial layer was finally infiltrated with Araldite and polymerised at 60°C for 48 h.

Sectioning and toluidine blue staining

Semi-thin survey sections (2 µm) were taken using a diamond knife and mounted onto glass slides, and the epithelial layer stained with Toluidine blue (TB). TB staining was performed by drying the sections; when the slides were dry, several drops of staining solution were placed onto the sections for 1–2 min depending on depth of colour required. Excess stain was removed by gently rinsing the slide with distilled water, the slides were then air dried and mounted in DPX.

Protein digestion and peptide labelling with iTRAQ reagents

Frozen cell pellets were resuspended in 10 mM triethylammonium bicarbonate with 1% (v/v) Nonidet P40 (NP®-40) and 10 mM phenylmethylsulfonyl fluoride (PMSF) and incubated for 15 min on ice. Lysates were clarified by centrifugation (10,000*g*, 5 min) and quantified using the Bradford assay (Bradford, 1976). Equal amounts of proteins (10 µg) were digested and labeled with iTRAQ reagent using the method and reagents supplied in the Applied Biosystems iTRAQ reagent multiplex kit. Briefly, the proteins were reduced with 5 mM tris (2-carboxyethyl) phosphine (TCEP) at 60°C for 60 min. Post-incubation the samples were alkylated with 10 mM methyl methanethiosulfonate (MMTS) for 10 min. Proteins were digested with 4 µg of porcine trypsin (Promega, UK) and incubated at 37°C overnight. One iTRAQ reagent vial was added to each sample and incubated at room temperature for 60 min (Brennan et al., 2009).

LC-MALDI

Equal volumes of the digested peptides labeled with different iTRAQ reagents were mixed, and peptides corresponding to 2 μg of undigested protein were separated on a nano-LC system (Ultimate 3000, Diones, Sunnyvale, USA) using a two-dimensional (2-D) salt plug method (Brennan et al., 2009). MS was performed using a MALDI TOF/TOF mass spectrometer (Applied Biosystems 4800 MALDI TOF/TOF Analyzer) as previously described (Brennan et al., 2009).

Data and statistical analysis

The MS/MS data was used to search the Swiss-Prot database (version 57.3; 468851 sequences; human taxonomy) using the MASCOT Database search engine version 2.1.04 (Matrix Science Ltd., London, UK), embedded into GPS-Explorer software version 3.6 Build 327 (Applied Biosystems; default GPS parameters, 1 missed cleavage allowed, fixed modification of MMTS(C), variable modifications of oxidation (M), pyro-glu (N-term E) and pyro-glu (N-term Q), 150 ppm mass tolerance in MS and 0.3 Da mass tolerance for MS/MS which are recommended published tolerances for LC-MALDI; Brennan et al., 2009.

Fold-changes (FC) between the treated samples and the reference control samples were calculated by the software for each of the quantified peptides of the same protein. Search results were evaluated by protein and peptide score confidence intervals % (CI%) were calculated in the GPS-Explorer software and based on the MASCOT score (Gluckmann et al., 2007). In order for a protein to be accepted there needed to be a minimum of two peptides with MASCOT e-values less than 0.05. Protein identification results with CI% score above 95 were accepted, and all the identifications in Table 1 were additionally manually inspected for correctness. Data from three independent biological replicates were examined for consistent protein changes that were altered by more than 1.5-fold.

1D SDS-PAGE, western blotting and antibody detection

Following protein quantification, an equal volume of 2x gel sample buffer (2x GSB: 0.1 M Tris HCl pH 6.8, 0.2 M DTT, 4% SDS, 20% glycerol, 0.1% bromophenol blue) was added and samples heated at 100°C for 5 min. Proteins were then separated by SDS-PAGE and transferred onto polyvinylidenedifluoride(PVDF)membrane(Amersham) before incubation with primary antibodies (Actin Sigma; A-2066 and Cystatin-A Santa Cruz; sc-32803) overnight at 4°C. Post-incubation, blots were washed three times and then incubated for 1 h with an appropriate horseradish peroxidise (HRP)-conjugated secondary antibody (Goat Anti-Rabbit IgG, HRP-linked Antibody Cell Signaling; 7074 and Goat Anti-Mouse IgG, HRP-linked Antibody Cell Signaling; 7076). Following this, the blots were washed and then incubated with Immobilon Western HRP Substrate (Millipore) for 5 min and then exposed to autoradiograph film. Following exposure each section of the blot was stripped and incubated in a sealed polyethylene bag with approximately 15–20 ml of MESNA stripping buffer for 30 min at 50°C.

Results

Quantitative proteomics

Quantitative proteomics was performed following treatments of the EpiAirway™ cells with the surrogate vapour and particulate phase components of tobacco smoke. A broad range of classes of protein were identified. Utilising iTRAQ technology, three technical replicates were generated and samples were lysed with 1% NP40 in a buffer. Protein was assayed and equal amounts were digested with trypsin and labeled with the iTRAQ reagents. The peptide mixtures were separated by nano-LC and eluted peptides were mixed with matrix and placed on a MALDI target plate. Protein identification and quantification was achieved by analysis of peptides by MALDI TOF/TOF mass spectrometry.

Seven different runs with different mixtures of the biological replicates were performed. Protein identifications were generated using the ABI Global Proteome System with MASCOT embedded in it. Using a threshold of a Total Ion Score of greater than forty, 466 different proteins were identified across all the runs with an average of 276 proteins identified in each run (see supporting information, Excel Sheet). These were made up of a wide variety of gene ontology groups when analysed for molecular function or biological process using GeneCoDis. Within the biological process ontology, the largest group of proteins was entitled "cellular biopolymer metabolic process" (GO:0034960) which described 154 protein IDs. For molecular function, 276 proteins were within the "protein binding" gene ontology group (GO:0005515). The top ten of both analyses are illustrated in Figure 1.

The data was analysed to determine which treatment caused the most changes outside of the 95% confidence interval for that particular treatment. The dataset for each treatment was used to calculate the Z-score and the number of samples with a Z-score greater than 1.96 or less than −1.96 was determined. The average number of protein changes across all the treatments and experiments was 11 proteins with 10 proteins increasing and only 1 protein decreasing in expression. None of the four TSC treatments altered significantly more proteins than another treatment, demonstrating a common magnitude of the numbers of proteins altered in response to the distinct agents in tobacco smoke.

Given the similar number of proteins altered by the different treatments, we investigated whether the overlap in the identity of the proteins altered, with a view to identifying common molecules that were "biomarkers" of the changes induced by tobacco smoke components. As very few proteins were down-regulated, we focused on proteins that were increased in expression. The average value of relative protein expression was calculated

Table 1. Overview of the proteomic expression data from exposed cell cultures. Proteins that were identified by parameters set out in the methods. Protein names and abbreviations from the Swiss-Prot database are indicated.

Protein Name	Accession No	Ave Pep Cnt	Ave Ion Score	Total Pep Count	Nicotine		Cadmium		Formaldehyde		Urethane	
					Rel Exp	*p*-val	Rel Exp	*p*-val	Rel Exp	*p*-val	Rel Exp	*p*-val
Aldo-keto reductase family 1 member C1	AK1C1_HUMAN	6	**298**	**19**	**1.30**		1.79		0.99		**1.53**	<0.05
Beta-2-microglobulin	B2MG_HUMAN	1	**87**	**7**	1.86	<0.05	**1.21**		**1.04**		**0.91**	
Calmodulin	CALM_HUMAN	4	**234**	**21**	**1.74**		**2.20**	<0.05	1.77	<0.05	1.75	<0.01
Dipeptidyl-peptidase 1	CATC_HUMAN	2	**135**	**12**	1.56		1.57		**0.95**		**1.55**	<0.05
CD9 antigen	CD9_HUMAN	2	**94**	**10**	**2.25**	<0.01	**2.44**	<0.01	**2.29**	<0.01	**1.72**	<0.01
Complement C3	C03_HUMAN	4	178	**21**	**2.04**	<0.01	1.69		**0.91**		**1.20**	
Cytochrome c oxidase subunit 7C, mitochondrial	COX7C_HUMAN	1	**62**	**4**	1.56		**2.11**	<0.05	**1.41**		**1.52**	<0.05
Cystatin-A	CYTA_HUMAN	3	**192**	**17**	**3.40**	<0.01	3.69	<0.01	**3.04**	<0.01	**3.22**	<0.01
Filamin-A	FLNA_HUMAN	1	**51**	**4**	**1.26**		**2.59**	<0.01	**1.31**		0.00	
ATP-dependent DNA helicase 2 subunit 2	KU86_HUMAN	1	**98**	**4**	1.78	<0.05	0.75		1.67		0.00	
Myristoylated alanine-rich C-kinase substrate	MARCS_HUMAN	2	**102**	**8**	**1.93**	<0.05	**1.03**		1.85	<0.05	**0.84**	
Histone H1.2	H12_HUMAN	2	**132**	**7**	**0.53**		0.87		**0.21**		**1.91**	<0.01
Myb-related protein A	MYBA_HUMAN	2	**57**	**3**	1.69		**2.54**	<0.01	**1.27**		**1.53**	<0.05
Neutrophil gelatinase-associated lipocalin	NGAL_HUMAN	5	**347**	**27**	1.87	<0.05	**1.51**		**1.19**		1.36	
Polymeric immunoglobulin receptor	PIGR_HUMAN	3	**142**	**16**	**2.61**	<0.01	1.88		**1.19**		**1.23**	
Protein Plunc	PLUNC_HUMAN	3	**184**	**18**	**2.07**	<0.01	**1.44**		1.68		**1.09**	
Cytochrome b-d complex subunit 7	QCR7_HUMAN	1	**81**	**5**	**2.15**	<0.01	**1.42**		**2.42**	<0.01	**0.94**	
Heterogeneous nuclear ribonucleoprotein AO	ROA0_HUMAN	1	**59**	**4**	**2.00**	<0.01	**3.53**	<0.01	**2.34**	<0.01	**2.25**	<0.01
Protein S100-A2	S10A2_HUMAN	5	**285**	**31**	1.63		**1.13**		1.73	0.05	**1.29**	
Cornifin-B	SPR1B_HUMAN	2	**93**	**9**	**2.22**	<0.01	0.71		**4.21**	<0.01	0.80	
Putative tropomyosin alpha-3 chain-like protein	TPM3L_HUMAN	1	**61**	**3**	1.96	<0.01	1.08		1.99	<0.01	0.81	
Ubiquitin	UBIQ_HUMAN	6	**355**	33	**2.37**	<0.01	**2.33**	<0.01	**2.81**	<0.01	**2.03**	<0.01
Cytochrome b-d complex subunit Rieske, mitochondrial	UCRI HUMAN	2	**84**	**5**	**1.30**		**2.61**	<0.01	**0.95**		**N/D**	

for each treatment from both the biological and technical replicates. A Z-score was calculated for each of these using the mean and standard deviation of each treatment. Only proteins that were identified in two or all of the biological replicates were used. A list was generated of the 23 proteins that were altered by one or more TSC. The average relative expression for these was calculated and the p-value was determined using the Z-score. This list is shown in Table 1 and the supporting information Excel Sheet. The average peptide count, ion score and the total peptide count was determined for each peptide ID. Our analysis revealed 4 proteins that were altered by all four TSC treatments: CD9 antigen, cystatin-A, heterogeneous nuclear ribonucleoprotein A0, and ubiquitin. We also identified three proteins associated with atherosclerosis, a disease associated with tobacco smoking. Furthermore, each treatment also caused unique changes in protein expression. We confirmed the expression of cystatin-A was increased by immunoblotting (Table 2).

Histopathology

Alterations in lung biochemistry and expression profiling are linked in biological context to histopathology. In order to relate the protein changes to a clinical end-point, semi-thin survey sections of resin-embedded EpiAirway™ samples, from all TSC exposures (versus controls), were stained with Toluidine Blue to enable visual confirmation of morphological changes in the epithelial layer architecture (Figures 2–6).

Cadmium treatment with 80 μM yielded focal regions of damage denoted by chromatin condensation and cytosolic retractions. A concentration of 160 μM induced

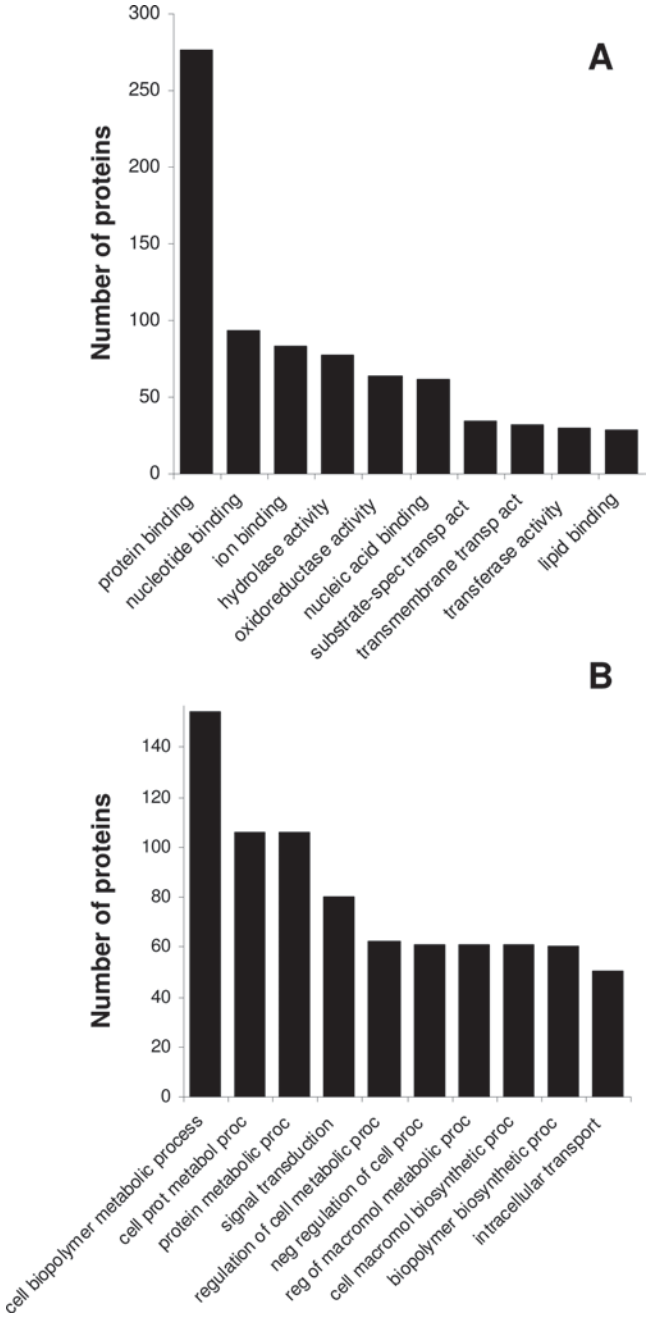

Figure 1. Gene ontology data for all tobacco smoke components. Genecodis was used to obtain the gene ontology for all proteins that were identified. All proteins, regardless of fold-change, were used in this analysis to identify: A) Biological processes and B) Molecular functions.

Table 2. Quantification of increased expression of cystatin-A by immunoblotting. Western blots were scanned and the bands quantified by densitometry using Genetools software ($n=3$). The cystatin-A induction result concurs with the data from the iTRAQ, experiment.

TSC	Fold increase of cystatin-A expression
Cadmium	4.2
Formaldehyde	3.1
Nicotine	2.5
Urethane	4.0

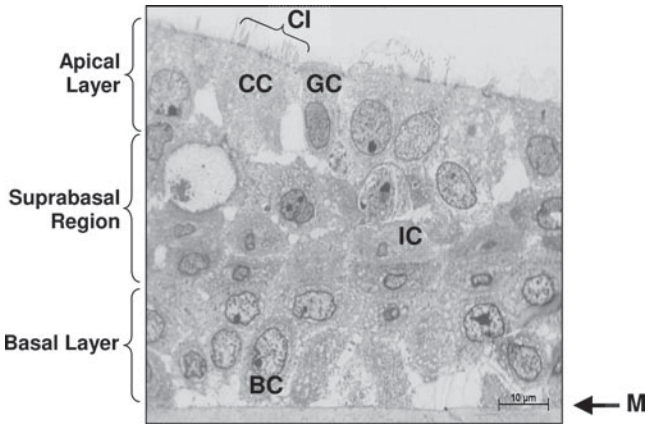

Figure 2. Semi-thin (2 µm) section of ELM stained with toluidine blue. ELM exposed to PBS maintained the pseudo-stratified, highly differentiated multi-layered model supported on a collagen membrane (M). Toluidine blue staining confirmed the presence of basal cells (BC), intermediate cells (IC), ciliated cells (CC), Goblet cells (GC), cilia on the apical surface (CI).

basal cell hyperplasia (BCH) and further decomposition of the cellular integrity. At 240 µM of cadmium, BCH was markedly increased. The epithelial layer exhibited a non-viable architecture (disrupted cellular junctions) at 320 µM (Figure 3). A 1 mM concentration of formaldehyde created regional hypertrophy (RH) with evidence of changes in cellular morphology/swelling (Figure 4). A concentration of 5 mM initiated early degradation of the epithelial layer along with reactive cell synthesis, depicted by a larger nucleus and expansion of cytoplasm (Figure 4). Further increased challenges caused changes in early cellular development, cell swelling and disruption of the plasma membrane. At the 10 mM dose, there was evidence of vacuolations leading to cytoplasmic blebbing, thinning of the epithelial layer and eventual sloughing-off of the cells (Figure 4). A 25 mM concentration of nicotine induced cytolysis of the cell components and a loss of cellular organisation. Epithelial layer degeneration was denoted by disruption of the intracellular junctions, presence of vacuoles and disordered epithelial layer architecture. This form of decomposition continued with further increments of nicotine treatments affecting diffuse lipid vacuolation and thickening of membranes (Figure 5). Diffuse cellular swelling and regional epithelial layer degeneration occurred with a 100 mM urethane concentration. Increased concentrations generated region-specific (basal region only) alterations in cells. Quantitative observations of the epithelial layer histopathology suggested that the nucleus-to-cytoplasmic ratios were enlarged when compared to ratios at the lower urethane concentrations (Figure 6).

Discussion

A major factor driving the increased burden of lung disease, especially COPD, is the human habit of smoking tobacco. Tobacco smoking, and the development of related pulmonary disease in humans, has been

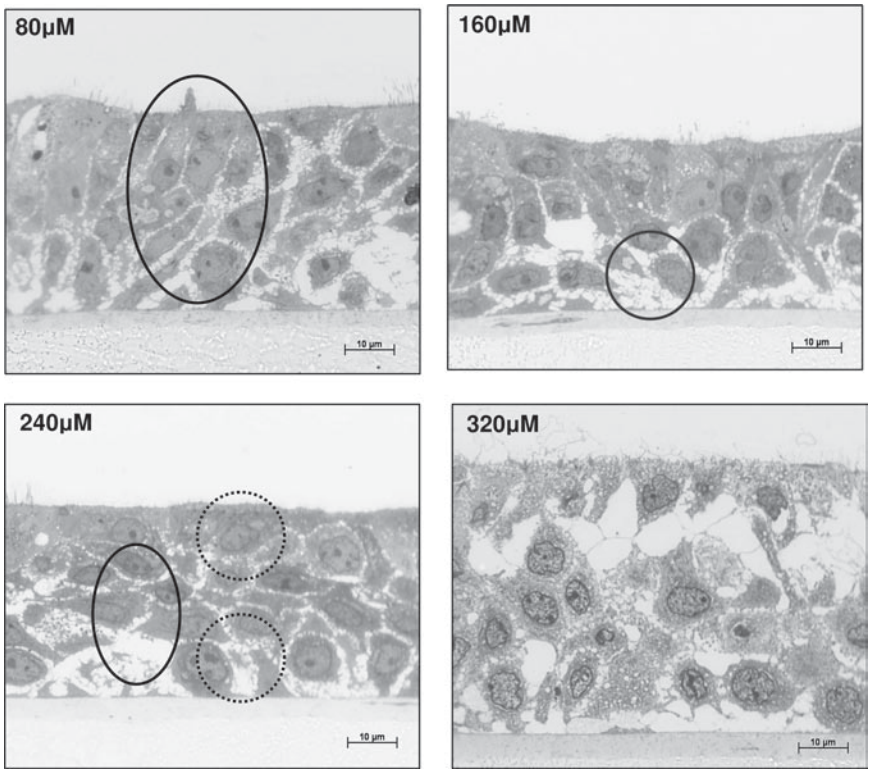

Figure 3. Semi-thin (2 μm) section of ELM stained with toluidine blue treated with increasing concentrations of cadmium. 80 μM: Black circle—focal regions of damage observed (chromatin condensation and cytosolic retractions); 160 μM: grey circle—basal cell hyperplasia; 240 μM: black circle—area of necrosis, dashed circle—areas of hyperplasia; 320 μM: complete loss viable architecture. scale bar 10 μm.

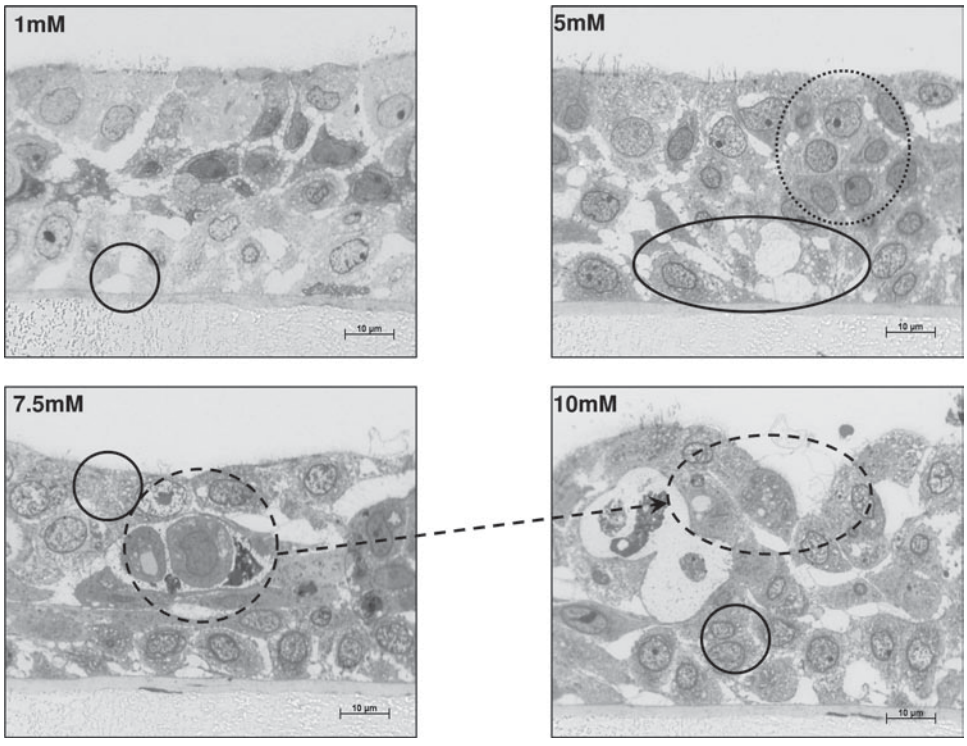

Figure 4. Semi-thin (2 μm) section of ELM stained with toluidine blue treated with increasing concentrations of formaldehyde. 1 mM: Black circle—area of regional hypertrophy; 5 mM: dashed circle—area of reactive cell synthesis, black circle—area epithelial layer degradation; 7.5 mM: Black circle—changes in early cellular development, cell swelling and disruption of the plasma membrane leading to early necrosis, Dashed circles—progression of cytoplasmic blebbing with increased dose; 10 mM: Black circle—vacuolation. Scale bar 10 μm.

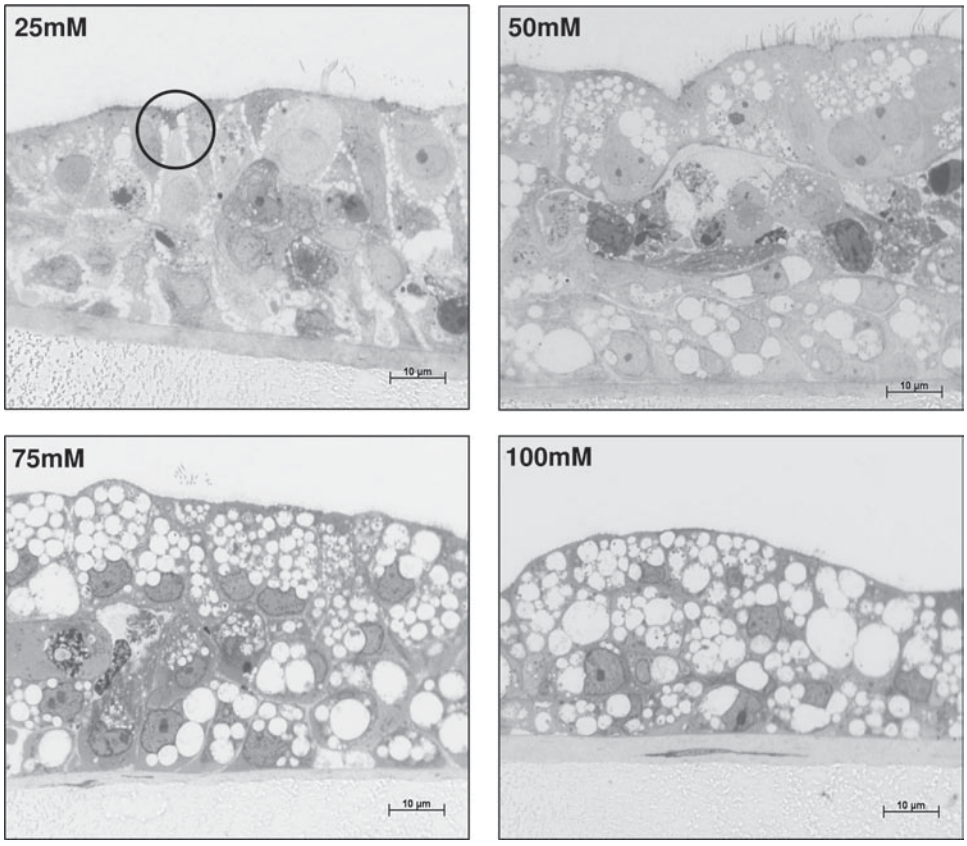

Figure 5. Semi-thin (2 µm) section of ELM stained with toluidine blue treated with increasing concentrations of nicotine. 25 mM: Black circle—(evidence of burst cells due to an osmotic imbalance) cytolysis further increases in nicotine induced disruption of the intracellular junctions and diffuse lipid vacuolation leading to thickening of membranes. Scale bar 10 µm.

historically assessed by epidemiological studies (Sexton et al., 2008). However, the number of chemicals used in tobacco approximately 4700; (Roberts, 1988; Green and Rodgman, 1996), and an individual's smoking habits i.e. different volumes of inhalation (Gowadia et al., 2009) confound these studies. In order to provide better resolution of the mechanisms of tobacco smoke toxicity, compounds with known mechanisms of action (i.e. thrombogenesis, carcinogenesis, mutagenesis and cytotoxicity), were selected from the particle phase (e.g. cadmium and nicotine) and vapour phase (e.g. formaldehyde and urethane) of tobacco smoke (TS). Recent advances in mass spectrometry have helped complement traditional protein-chemistry methods in detection and identification of proteins. New proteomic strategies based on isotope-coded affinity tags can accurately quantify individual proteins in complex mixtures, a capability that has been lacking in the past, and can also provide sequence information linking descriptive and quantitative proteomics (Gygi and Aebersold, 2000; Karp et al., 2010). Proteomic analysis of TSC from the vapour and particle phases may be useful in establishing biomarkers indicative of injury and disease patterns that are characteristic for a given constituent of tobacco smoke. This holds great potential for diagnostic, prognostic, and therapeutic applications through integration of basic and clinical data in a translational fashion.

Evolving proteomic technologies have the potential to identify each protein expressed in isolated cells or in complex tissues (e.g. EpiAirway™), and also to determine post-translational modifications, protein–protein interactions, and other critical features that define phenotype and regulate cellular function (Gygi and Aebersold, 2000). Proteomic approaches provide physiologic relevance that is not inherent in measuring transcript levels and profiles (Pradet-Balade et al., 2001). The proteome reflects both the intrinsic genetic programme of the cell and the impact of its immediate environment, and is therefore valuable in biomarker discovery. Post-translational modifications of proteins, not detected through RNA analysis, may occur at different stages of lung injury/disease to provide fingerprints or intelligent biomarkers of exposure and harm to a toxicant (Matthay et al., 2003). The analysis of these biomarker patterns provides a means of classifying injury and disease heterogeneity, whereby specific proteomic profiles can be correlated with mechanistic qualities of tobacco smoke components. In addition, the most effective use of proteomic strategies, as with genomic technology e.g. Sexton et al., 2008, is in combination with conventional toxicology and pathology to determine the functions of key proteins and their effects on cell and tissue behavior.

Advances in 'omic technologies are also dramatically increasing the need to evaluate large numbers of

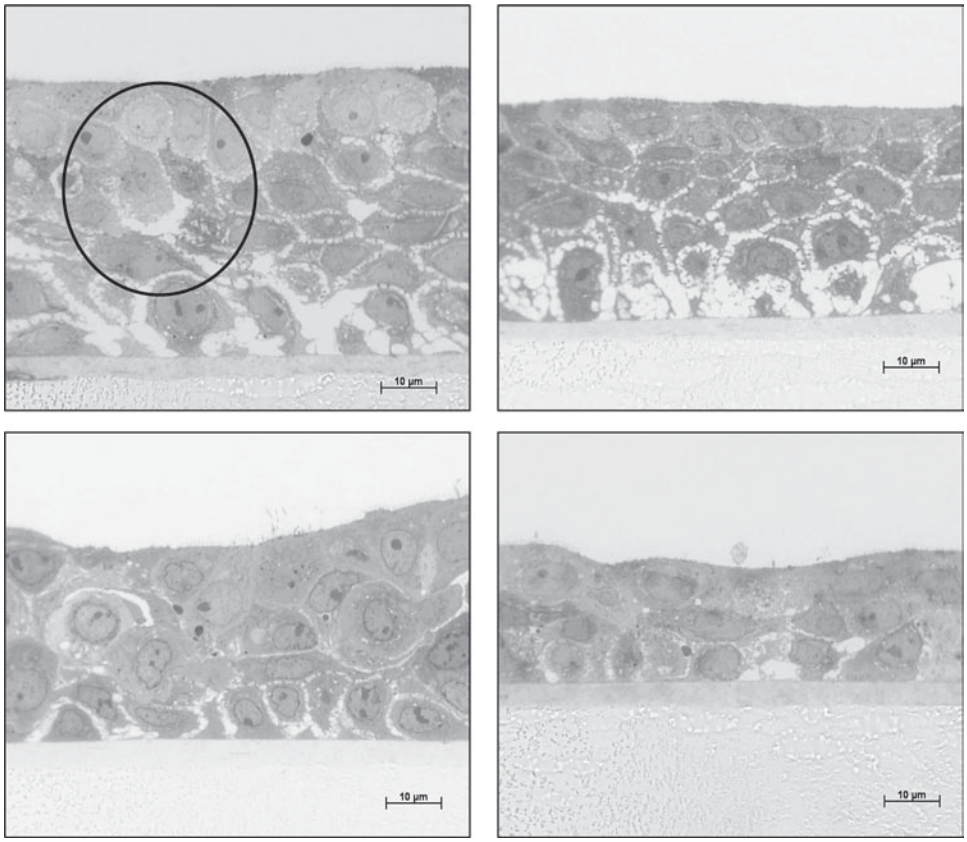

Figure 6. Semi-thin (2 µm) section of ELM stained with toluidine blue treated with increasing concentrations of urethane. 250 mM: Black circle—indicates diffuse cellular swelling with regional hypertrophy; 500 mM: Changes in basal cellular development characterised by expansion of the cytoplasm and enlargement of the cell nucleus. Scale bar 10 µm.

molecular targets for their diagnostic, predictive or prognostic value. Conventional toxicology techniques are often tedious, time-consuming, and require considerable amounts of tissue (animal or human), thereby limiting both the number of tissues and the number of targets that can be evaluated. Here, we demonstrate the power of our previously described work with the EpiAirway™ lung construct e.g. (Balharry et al., 2008; Sexton et al., 2008), in conjunction with a combination of iTRAQ, nano-LC and MALDI TOF/TOF technologies, that overcomes these caveats. This work represents the first quantitative proteomic investigation of the effects of TSC on a complex airway epithelial model, used to replicate *in vivo*-like exposures, in the human lung.

There have been two recent studies that have utilised human lung samples to investigate the proteomic profiles of smokers compared to non-smokers (Lee et al., 2009; Steiling et al., 2009), the latter concentrating on COPD. In particular, similarities have been shown between these studies, and this is at least in part due to the range of complex techniques that can be employed for proteomic profiling. However, the ontology results between one of these studies (Steiling et al., 2009) and this current investigation do show some similarities, in that erythrocyte contamination was evident in the primary tissue; a common complication of harvesting cells from a living donor.

Our study demonstrated that the proteomic analysis of the EpiAirway™ multi-cellular *in vitro* lung model can be successful and can provide an alternative to animal testing. The morphological changes induced by TSC were similar to the changes shown to occur *in vivo* (Innes et al., 2006) and many of the molecules identified by this investigation have a relevance to lung biology. Through this study, we have identified both shared and unique changes for the 4 types of TSC investigated. This epithelial layer model presented three important advantages. First, the use of an *in vitro* model allowed a reduction in animal numbers required for *in vivo* studies. Importantly, some of the lung changes observed *in vitro* mimicked those seen *in vivo*, demonstrating the validity of the model (Balharry et al., 2008). Second, the *in vitro* model was composed of human cells, meaning that there was a direct relevance to pathogenic changes seen following smoking. This also allowed comparisons with pathogenic samples from human subjects. From a bioinformatics point-of-view, the study of human tissue was advantageous because the human protein database was well-annotated. The third advantage was that it permitted a detailed investigation of lung pathology, one of the most important tissues altered during smoking (Caminati et al., 2010).

Because of these advantages, it was critical to investigate this model using the current technology of quantitative proteomics. A unique facet of this study was the use of a lung model that was comprised of different cell types. In this way, it reflected the lung better than a single

cell line. Furthermore, each of the cell types was primary human tissue and not an immortalised malignant cell line, again making the comparison to *in vivo* lung more relevant. The quality of the data and the number of protein identifications achieved was similar to a study on human bronchial airway epithelium (Steiling et al., 2009), demonstrating that the proteomic approach can be applied to a multi-cell type *in vitro* model.

One of the most interesting parts of this study was the proteins that were shared by all 4 of the TSC. These molecules were cystatin-A, CD9 antigen, heterogeneous nuclear ribonucleoprotein A0 (ROA0) and ubiquitin. Interestingly, these proteins also showed the largest magnitude change, albeit not in the cases of all the treatments. Examples included: ROA0 which had a 3.5-fold increase in response to cadmium; and cystatin-A which had a greater than 3-fold change in response to all TSC. Because it showed the largest change in expression, cystatin-A expression was also measured with specific antibodies, and was shown to be increased by each treatment. There was evidence for a role for three of these proteins in cancer. Cystatin-A has been shown to be elevated in lung cancer (Leinonen et al., 2007) and cystatin function has been proposed as a novel cancer therapeutic target. Thus, the increase of cystatin, in particular, may have relevance both as a biomarker and as a therapeutic (Keppler, 2006). CD9 expression has been measured in breast cancer and non-small cell lung cancer. However, decreased expression indicated poor prognosis, perhaps due to the role of CD9 in adhesion (Higashiyama et al., 1995; Miyake et al., 1996; Funakoshi et al., 2003). Mouse studies suggested CD9 played a role in lung inflammation; therefore the increase in our model may be an inflammatory rather than malignant change (Suzuki et al., 2009). Ubiquitin ligases, which utilise ubiquitin have altered expression in cancer (Confalonieri et al., 2009). Thus, up-regulation of their substrate, ubiquitin, would be expected to be necessary for function of these ligases.

In contrast, ROA0, has not been studied in cancer, but has been shown to regulate cytokine gene expression in inflammatory conditions (Rousseau et al., 2002). Interestingly, we have previously identified the pro-inflammatory cytokines IL-1 and IL-6 as being elevated in this lung model by the same four TSC (Balharry et al., 2008). ROA0 has been demonstrated to be involved in the regulation of IL-6 mRNA stability. Thus, an increase in ROA0 may be a molecular mechanism that contributes to inflammation of the lung by TSC.

Tobacco smoking initiates a number of mechanisms such as thrombosis and vascular inflammation predisposing to atherosclerosis (Lee and Cooke, 2011). Nicotine is a compound that has been identified as contributing to these atherogenesis factors (Benowitz, 2003). This study identified three proteins associated with atherosclerosis (complement C3, calmodulin and CD9). Studies have shown that atherosclerosis is an inflammatory disease and the protein C3 is a mediator in the inflammatory process (Szeplaki et al., 2004). Another aspect of the disease is the presence of atheromatous plaques (Gamble, 2006); induction of the protein calmodulin up-regulates plaque precursor matrix which initiates this disease process. CD9 is a cell surface protein and has been implicated in platelet activation and aggregation events (Jennings et al., 1990). CD9 has also been shown to enhance the juxtacrine mitogenic activity, which in turn, can lead to the development of atherosclerosis (Ouchi et al., 1997).

The majority of changes were unique to each of the TSC: 11 proteins were only significantly altered by one of the TSC. While our data did not suggest that any single component was more active than others, in general, there would appear to be pathogenic changes that were unique to each treatment. Interestingly, nicotine had the most unique changes, 6 proteins altered as compared to 3 or less in each of the other components. Nicotine also altered a distinct set of genes in our gene expression analysis of these components in this epithelial layer model (Sexton et al., 2008). As we only measured a subset of genes, we saw no overlap with our proteomic data but both data sets suggested that nicotine functions differently to the other TSC.

Conclusions

In summary, we report a quantitative proteomic analysis of an EpiAirway epithelial layer model upon the exposure to four different tobacco smoke components: nicotine, cadmium, formaldehyde and urethane. Four proteins were altered by all TSC components and we validated the most increased protein: cystatin-A. Three of these proteins have been implicated in lung cancer while the third was involved in regulating inflammatory gene expression. These may be useful markers of tissue damage in the lung.

Epidemiological studies have clearly established that tobacco products cause a variety of diseases, especially in the lung and heart, and improving our understanding of the relevant molecular mechanisms will lead to new approaches for their prevention. The challenges of discovering the mechanisms behind cigarette smoking-related diseases remains substantial, and modern evolving approaches in fundamental science, such as proteomics, provides substantial opportunities to improve our understanding. Integration of molecular biomarkers into epidemiological studies could provide crucial mechanistic information about the effects of exposures (intentional or unintentional), leading to new approaches for disease detection and prevention.

Acknowledgement

This work was funded by the Institute for Science and Health, USA. The proteomics was carried by Central Biotechnology Services (CBS), Cardiff University. Special thanks to Dr Martin Foster for his help with the pathological interpretation.

Declaration of interest

There is no conflict of interests for any of the authors.

References

Balharry D, Sexton K, BéruBé KA. (2008). An *in vitro* approach to assess the toxicity of inhaled tobacco smoke components: nicotine, cadmium, formaldehyde and urethane. *Toxicology* 244:66–76.

Benowitz NL. (2003). Cigarette smoking and cardiovascular disease: pathophysiology and implications for treatment. *Prog Cardiovasc Dis* 46:91–111.

Bérubé K, Prytherch Z, Job C, Hughes T. (2010). Human primary bronchial lung cell constructs: the new respiratory models. *Toxicology* 278:311–318.

BéruBé K, Aufderheide M, Breheny D, Clothier R, Combes R, Duffin R, Forbes B, Gaça M, Gray A, Hall I, Kelly M, Lethem M, Liebsch M, Merolla L, Morin JP, Seagrave J, Swartz MA, Tetley TD, Umachandran M. (2009). *In vitro* models of inhalation toxicity and disease. The report of a FRAME workshop. *Altern Lab Anim* 37:89–141.

Bérubé KA, Balharry D, Jones TP, Moreno T, Hayden P, Sexton K, Hicks M, Merolla L, Timblin C, Shukla A, Mossman BT. (2006). Characterisation of airborne particulate matter and related mechanisms of toxicity: an experimental approach. *In*: J., Ayres, R., Maynard & Richards, R (eds.) *Air Pollution Reviews*. London: Imperial College Press.

Bradford MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254.

Brennan P, Shore AM, Clement M, Hewamana S, Jones CM, Giles P, Fegan C, Pepper C, Brewis IA. (2009). Quantitative nuclear proteomics reveals new phenotypes altered in lymphoblastoid cells. *Proteomics Clin Appl* 3:359–369.

Caminati A, Graziano P, Sverzellati N, Harari S. (2010). Smoking-related interstitial lung diseases. *Pathologica* 102:525–536.

Confalonieri S, Quarto M, Goisis G, Nuciforo P, Donzelli M, Jodice G, Pelosi G, Viale G, Pece S, Di Fiore PP. (2009). Alterations of ubiquitin ligases in human cancer and their association with the natural history of the tumor. *Oncogene* 28:2959–2968.

Funakoshi T, Tachibana I, Hoshida Y, Kimura H, Takeda Y, Kijima T, Nishino K, Goto H, Yoneda T, Kumagai T, Osaki T, Hayashi S, Aozasa K, Kawase I. (2003). Expression of tetraspanins in human lung cancer cells: frequent downregulation of CD9 and its contribution to cell motility in small cell lung cancer. *Oncogene* 22:674–687.

Gamble W. (2006). Atherosclerosis: the carbonic anhydrase, carbon dioxide, calcium concerted theory. *J Theor Biol* 239:16–21.

Glauert A, Lewis P. (1998). Biological specimen preparation for transmission electron microscopy. Chapter 5. Embedding methods. *In*: Glauert, A (ed.) *Practical Methods in Electron Microscopy*.

Glückmann M, Fella K, Waidelich D, Merkel D, Kruft V, Kramer PJ, Walter Y, Hellmann J, Karas M, Kröger M. (2007). Prevalidation of potential protein biomarkers in toxicology using iTRAQ reagent technology. *Proteomics* 7:1564–1574.

Gowadia N, Oldham MJ, Dunn-Rankin D. (2009). Particle size distribution of nicotine in mainstream smoke from 2R4F, Marlboro Medium, and Quest1 cigarettes under different puffing regimens. *Inhal Toxicol* 21:435–446.

Green CR, Rodgman A. (1996). The tobacco chemists' research conference; A half century of advances in analytical methodology of tobacco and its products. *Recent Adv Tob Sci* 22:131–304.

Gygi SP, Aebersold R. (2000). Mass spectrometry and proteomics. *Curr Opin Chem Biol* 4:489–494.

Higashiyama M, Taki T, Ieki Y, Adachi M, Huang CL, Koh T, Kodama K, Doi O, Miyake M. (1995). Reduced motility related protein-1 (MRP-1/CD9) gene expression as a factor of poor prognosis in non-small cell lung cancer. *Cancer Res* 55:6040–6044.

Hoffmann D, Hoffmann I, El-Bayoumy K. (2001). The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder. *Chem Res Toxicol* 14:767–790.

Innes AL, Woodruff PG, Ferrando RE, Donnelly S, Dolganov GM, Lazarus SC, Fahy JV. (2006). Epithelial mucin stores are increased in the large airways of smokers with airflow obstruction. *Chest* 130:1102–1108.

Jemal A, Siegel R, Xu J, Ward E. (2010). Cancer statistics, 2010. *ca Cancer J Clin* 60:277–300.

Jennings LK, Fox CF, Kouns WC, McKay CP, Ballou LR, Schultz HE. (1990). The activation of human platelets mediated by anti-human platelet p24/CD9 monoclonal antibodies. *J Biol Chem* 265:3815–3822.

Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS. (2010). Addressing accuracy and precision issues in iTRAQ quantitation. *Mol Cell Proteomics* 9:1885–1897.

Keppler D. (2006). Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett* 235:159–176.

Lancet. (2009). Tobacco smoking: why start? *Lancet* 374:1038.

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

Lee J, Cooke JP. (2011). The role of nicotine in the pathogenesis of atherosclerosis. *Atherosclerosis* 215:281–283.

Leinonen T, Pirinen R, Böhm J, Johansson R, Rinne A, Weber E, Kosma VM. (2007). Biological and prognostic role of acid cysteine proteinase inhibitor (ACPI, cystatin A) in non-small-cell lung cancer. *J Clin Pathol* 60:515–519.

Matthay MA, Zimmerman GA, Esmon C, Bhattacharya J, Coller B, Doerschuk CM, Floros J, Gimbrone MA Jr, Hoffman E, Hubmayr RD, Leppert M, Matalon S, Munford R, Parsons P, Slutsky AS, Tracey KJ, Ward P, Gail DB, Harabin AL. (2003). Future research directions in acute lung injury: summary of a National Heart, Lung, and Blood Institute working group. *Am J Respir Crit Care Med* 167:1027–1035.

Miyake M, Nakano K, Itoi SI, Koh T, Taki T. (1996). Motility-related protein-1 (MRP-1/CD9) reduction as a factor of poor prognosis in breast cancer. *Cancer Res* 56:1244–1249.

Ouchi N, Kihara S, Yamashita S, Higashiyama S, Nakagawa T, Shimomura I, Funahashi T, Kameda-Takemura K, Kawata S, Taniguchi N, Matsuzawa Y. (1997). Role of membrane-anchored heparin-binding epidermal growth factor-like growth factor and CD9 on macrophages. *Biochem J* 328 (Pt 3):923–928.

Pradet-Balade B, Boulme F, Beug H, Mullner E, Garcia Sanz J. (2001). Cytokine balance translation control: bridging the gap between genomics and proteo- in the lungs of patients with acute respiratory distress syndrome. *Trends Biochem Sci* 26:225-229.

Roberts DL. (1988). Natural tobacco flavour. *Adv Tob Sci* 14:49–81.

Rousseau S, Morrice N, Peggie M, Campbell DG, Gaestel M, Cohen P. (2002). Inhibition of SAPK2a/p38 prevents hnRNP A0 phosphorylation by MAPKAP-K2 and its interaction with cytokine mRNAs. *Embo J* 21:6505–6514.

Sexton K, Balharry D, BéruBé KA. (2008). Genomic biomarkers of pulmonary exposure to tobacco smoke components. *Pharmacogenet Genomics* 18:853–860.

Steiling K, Kadar AY, Bergerat A, Flanigon J, Sridhar S, Shah V, Ahmad QR, Brody JS, Lenburg ME, Steffen M, Spira A. (2009). Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. *Plos one* 4:e5043.

Suzuki M, Tachibana I, Takeda Y, He P, Minami S, Iwasaki T, Kida H, Goya S, Kijima T, Yoshida M, Kumagai T, Osaki T, Kawase I. (2009). Tetraspanin CD9 negatively regulates lipopolysaccharide-induced macrophage activation and lung inflammation. *J Immunol* 182: 6485–6493.

Széplaki G, Prohászka Z, Duba J, Rugonfalvi-Kiss S, Karádi I, Kókai M, Kramer J, Füst G, Kleiber M, Romics L, Varga L. (2004). Association of high serum concentration of the third component of complement (C3) with pre-existing severe coronary artery disease and new vascular events in women. *Atherosclerosis* 177:383–389.