J Mol Med (2006) 84: 318–328
DOI 10.1007/s00109-005-0008-2

ORIGINAL ARTICLE

Adriana Heguy · Timothy P. O'Connor · Karsta Luettich · Stefan Worgall · Adam Cieciuch · Ben-Gary Harvey · Neil R. Hackett · Ronald G. Crystal

Gene expression profiling of human alveolar macrophages of phenotypically normal smokers and nonsmokers reveals a previously unrecognized subset of genes modulated by cigarette smoking

Received: 17 January 2005 / Accepted: 21 September 2005 / Published online: 7 March 2006

Abstract Cigarette smoking is the leading cause of the respiratory diseases collectively known as chronic obstructive pulmonary disease (COPD). While the pathogenesis of COPD is complex, there is abundant evidence that alveolar macrophages (AM) play an important role. Based on the concept that COPD is a slow-progressing disorder likely involving multiple mediators released by AM activated by cigarette smoke, the present study focuses on the identification of previously unrecognized genes that may be linked to early events in the molecular pathogenesis of COPD, as opposed to factors associated with the presence of disease. To accomplish this, microarray analysis using Affymetrix microarrays was used to carry out an unbiased survey of the differences in gene expression profiles in the AM of phenotypically normal, ~20 pack-year smokers compared to healthy nonsmokers. Although smoking did not alter the global gene expression pattern of AM, 75 genes were modulated by smoking, with 40 genes up-regulated and 35 down-regulated in the AM of smokers compared to

Adriana Heguy received her PhD from the University of Southern California, Los Angeles, in the field of regulation of gene expression. She is currently an Assistant Professor in the Department of Genetic Medicine at the Weill Medical College of Cornell University. Her primary research interest is gene expression profiling in the normal and diseased human lung.

Ronald G. Crystal obtained his MD degree from the University of Pennsylvania. Dr. Crystal is Professor and Chairman of the Department of Genetic Medicine of the Weill Medical College of Cornell University, where he is also the Bruce Webster Professor of Intenal Medicine, Director of the Belfer Gene Theraphy Core Facility and Chief of the Division of Pulmonary and Critical Care Medicine at the Weill Cornell-New York Presbyterian Hospital. His research interests are in developing genetic medicines based on understanding gene expression in health and disease.

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00109-005-0008-2.

A. Heguy · T. P. O'Connor · K. Luettich · S. Worgall · R. G. Crystal
Department of Genetic Medicine,
Weill Medical College of Cornell University,
515 East 71st Street, S-1000,
New York, NY 10021, USA

S. Worgall
Department of Pediatrics,
Weill Medical College of Cornell University,
New York, NY, USA

B.-G. Harvey · R. G. Crystal
Division of Pulmonary and Critical Care Medicine,
Weill Medical College of Cornell University,
New York, NY, USA

A. Cieciuch · N. R. Hackett · R. G. Crystal (✉)
Belfer Gene Therapy Core Facility,
New York, NY, USA
e-mail: geneticmedicine@med.cornell.edu
Tel.: +1-212-7462258
Fax: +1-212-7468383

nonsmokers. Most of these genes belong to the functional categories of immune/inflammatory response, cell adhesion and extracellular matrix, proteolysis and antiproteolysis, lysosomal function, antioxidant-related function, signal transduction, and regulation of transcription. Of these 75 genes, 69 have not been previously recognized to be up- or down-regulated in AM in association with smoking or COPD, including genes coding for proteins belonging to all of the above categories, and others belonging to various functional categories or of unknown function. These observations suggest that gene expression responses of AM associated with the stress of cigarette smoking are more complex than previously thought, and offer a variety of new insights into the complex pathogenesis of smoking-induced lung diseases.

Keywords Chronic obstructive pulmonary disease · Molecular pathogenesis · Immune response and inflammation

Introduction

Cigarette smoking is the leading risk factor for the development of chronic obstructive pulmonary disease (COPD), including chronic bronchitis and emphysema [1, 2]. For most individuals with COPD, the disease is caused by the chronic burden of oxidants, particles, and xenobiotics imposed by chronic cigarette smoking on the components of the lung [3, 4]. COPD is a slow, progressive disease that does not manifest until at least >20 pack-years of smoking [1, 2, 5–7]. Although the molecular mechanisms involved in the pathogenesis of COPD are very complex, there is considerable data suggesting that alveolar macrophages (AM) play a significant role in the pathogenesis of COPD associated with cigarette smoking [8, 9]. In this regard, individuals who smoke but are otherwise phenotypically normal have approximately three times more AM on their airways and alveolar surface [8–10]. Compatible with this observation, the earliest pathologic change in response to moderate smoking is a respiratory bronchiolitis associated with clusters of intraluminal AM [11, 12]. Exposure of the AM of smokers to tobacco smoke results in AM activation [8–10]. Consistent with these data, the AM of healthy smokers express elevated levels of a variety of mediators, including neutrophil chemoattractants and other immune response and inflammatory mediators, proteases, and cell adhesion molecules, among others [8, 13, 14]. The current concepts of the pathogenesis of COPD hold that the burden of these AM-derived mediators plays a major role in the lung damage that characterizes this disorder [2, 8, 9].

The focus of the present study is to further identify the AM-derived processes with potential relevance to the pathogenesis of COPD, based on the hypothesis that an unbiased assessment of early gene expression changes, taking place before the establishment of COPD, will help to identify processes not previously associated with tobacco smoking or COPD. The strategy is to employ microarray analysis to survey the differences in gene expression profiles of the AM of phenotypically normal, 20 pack-year smokers compared to normal nonsmokers. The assessment of ten individuals (five nonsmokers, five current smokers) revealed up-regulation or down-regulation among smokers of a specific subset of 75 genes (40 up-regulated and 35 down-regulated), in the functional categories of immune/inflammatory response, cell adhesion and extracellular matrix, proteolysis and antiproteolysis, lysosomal function, antioxidant-related function, signal transduction and regulation of transcription, and others. Of these 75 genes, 69 have not been previously linked with AM in association with smoking or COPD.

Materials and methods

Study subjects

This study was approved by the Weill Cornell Medical College Institutional Review Board. Written informed consent was obtained from each individual prior to enrollment in the study. Individuals underwent an initial screening evaluation including history (detailed smoking habits), complete physical exam, blood studies, urine analysis, chest roentgenogram, lung function tests, and electrocardiogram (EKG). Special screening evaluation relevant to smoking habits included the urinary levels of nicotine and its derivative cotinine, and serum levels of carboxyhemoglobin. Upon completion of the baseline evaluation, those individuals who met the inclusion criteria (five nonsmokers and five smokers) underwent fiber-optic bronchoscopy and bronchoalveolar lavage (BAL) to obtain AM.

Collection of alveolar macrophages

Fiber-optic bronchoscopy was performed to obtain cells present in the BAL fluid using methods developed in our laboratory to ensure high quality RNA for gene expression analysis [15]. The total volume used per site was typically 100 ml. A maximum of three sites were evaluated, with a total volume not exceeding 300 ml. Recovery of the infused volume was typically 45–65%. The right middle lobe, right lower lobe, and lingula were the usual sites for lavage. BAL fluid was filtered with gauze and centrifuged at 1,200 rpm for 5 min, 4°C. Cells were washed twice in RPMI 1640 containing 10% fetal bovine serum, 50 U/ml penicillin, 50 U/ml streptomycin and 2 mM glutamine (Invitrogen, Carlsbad, CA), suspended in 10 ml medium, and an aliquot of 0.5 ml was used for a differential cell count. Cell viability was estimated by Trypan blue exclusion and expressed as a percentage of the total cells recovered. Total cell number was determined by counting on a hemocytometer. Differential cell count was assessed on sedimented cells prepared by cytocentrifugation (Cytospin 3; Shandon Instruments, Pittsburgh, PA) stained with DiffQuik (Baxter Healthcare, Miami, FL). The remainder was processed for RNA extraction, by seeding the

cells in six-well plastic culture dishes (2×10^6 per 2 ml/well) and purifying the AM by 2 h adherence at 37°C in a 5% CO_2 humidified incubator, removing the nonadherent cells by washing with RPMI 1640.

RNA extraction and preparation for Affymetrix microarrays

Total RNA was extracted using the TRIzol (Life Technologies, Gaithersburg, MD) method followed by RNeasy clean-up (Qiagen, Valencia, CA) to remove residual DNA, a procedure giving a yield of 2 to 4 μg from 10^6 cells. Complementary DNA (cDNA) and complementary RNA (cRNA) synthesis was prepared, and hybridized to the Affymetrix GeneChip HuGeneFL microarray, which enables the relative monitoring of messenger RNA (mRNA) transcripts of approximately 5,600 full-length human genes (~6,800 probes), initially released by Affymetrix in November of 1998. All procedures were carried out as specified by Affymetrix (Santa Clara, CA).

Microarray data analysis

The data on each individual microarray chip was scaled to an arbitrary target intensity, as recommended by Affymetrix, using the Microarray Suite version 5.0 software. Normalization was carried out using the GeneSpring software (Agilent Biotechnologies, Palo Alto, CA) as follows: (1) per microarray sample, dividing the raw data by the 50th percentile of all measurements; and (2) per gene, by dividing the raw data by the median of the expression level for the gene in all samples. To eliminate those genes not expressed in the AM, only the genes with detectable expression in at least one out of the ten samples (Affymetrix Detection Call of Present in at least one of the ten samples) were chosen for further analysis. The statistical analysis was carried out for these 4,199 genes. Fold-changes were calculated as the ratio of the average expression level in the smokers to the average expression level in the nonsmokers.

Clustering and tree building programs were used to compare the overall gene expression patterns among samples from smokers and nonsmokers for both global comparisons of all 4,199 genes flagged as Present in at least one sample, as well as evaluations of the genes that were found to be differentially expressed in the smokers compared to the nonsmokers (see Statistics). Normalized, log-transformed gene expression levels were evaluated using the Cluster program [16] and subjected to hierarchical complete linkage clustering by both individual and gene. The resulting cluster was visualized with the TreeView program [16].

TaqMan mRNA analysis

To confirm the results of the microarray analysis, TaqMan real-time reverse transcriptase (RT) polymerase chain reaction (PCR) analysis was used as an independent method of measuring gene expression levels. Samples from all five nonsmokers and four of the five smokers were assessed for three genes representative of novel observations [osteopontin, a disintegrin and metalloprotease domain 10 (ADAM10), and chemokine (C-X-C motif) ligand 6]. First strand cDNA was synthesized from 2 μg of RNA in a 100 μl reaction volume, using the TaqMan Reverse Transcriptase Reaction Kit (Applied Biosystems, Foster City, CA), with random hexamers as primers, and diluted with Universal Master Mix (Applied Biosystems) to 1:100 or 1:10. The probe and primers specific for mRNA were designed for each gene using the PrimerExpress software (Applied Biosystems). Each dilution was assayed in triplicate wells. Relative expression levels were calculated using the ΔΔ Ct method (Applied Biosystems), with ribosomal RNA (rRNA) as the internal control (Human Ribosomal RNA Kit, Applied Biosystems), and a cocktail consisting of equal parts of mRNA samples from the AM of the nonsmokers in this study, as the calibrator. The rRNA probe was labeled with VIC, and the probe for each of the three specific genes was labeled with FAM. The PCR reactions were run in an Applied Biosystems Sequence Detection System 7700. The relative quantity was calculated using the algorithm provided by Applied Biosystems.

Statistics

Comparison of the age of the subjects, cell yield and viability, and % cells types in the smokers and nonsmokers was performed by a two-tailed Student's *t* test. The significance of gene expression differences between the two groups was determined by calculating the *p* value for expression levels between the nonsmoker group and the smoker group using the Student's *t* test, assuming a two-tailed distribution and equal variances, with the log of the signal to background ratio as the starting value, using the GeneSpring software. To compare the results obtained using microarrays to those obtained using TaqMan real-time RT-PCR, a two-way analysis of variance (ANOVA) was performed, using method (microarray vs TaqMan) and smoking status (smokers vs nonsmokers) as independent factors. For the ANOVA, expression levels were normalized separately for the microarray and TaqMan analysis by dividing individual values by the average expression level of all nonsmokers and smokers for that method, to allow direct comparisons of values between the two methods.

Results

Study population and alveolar macrophage samples

The study population included ten individuals (all men; five healthy nonsmokers and five phenotypically normal current smokers, Supplemental Table 1). The smokers had an average smoking history of 19±3 pack-years. The two groups were similar in regard to age ($p>0.7$, smokers'

average 33±7 years; nonsmokers' average 37±6 years). All individuals were classified as normal based on standard medical history, physical exam, routine blood and urine tests, chest roentgenogram, EKG, and pulmonary function tests. Urine nicotine, urine cotinine, and venous carboxyhemoglobin levels verified that the individuals who gave a history of current smoking were smokers and that those reporting nonsmoking were nonsmokers. Approximately three times as many macrophages were recovered from the BAL samples of smokers as compared to those recovered from the BAL samples nonsmokers ($p<0.003$, Supplemental Table 1). The purity and viability of the samples obtained from the non-smokers was comparable to the purity of the samples obtained from smokers ($p>0.6$), and neither group had significant numbers (1%) of polymorphonuclear cells contaminating the preparations.

Global gene expression patterns in the alveolar macrophages of healthy smokers and nonsmokers

Global gene expression patterns, i.e., the overall expression of 4,199 genes expressed in these samples, did not distinguish AM from smokers from those of nonsmokers. Clustering and tree building programs [16] were used to compare the overall gene expression pattern among all samples from smokers and nonsmokers, using these 4,199 genes (Fig. 1). The data show no overall clustering of global gene expression patterns in smokers as compared to nonsmokers.

Representative individual-to-individual comparisons (smoker to smoker, nonsmoker to nonsmoker and smoker to nonsmoker) of expression levels (normalized by array only) of the 1,582 genes present on all ten microarrays indicated overall similarity in global gene expression levels among individuals, because all comparisons (smokers to smokers, nonsmokers to nonsmokers and smokers to nonsmokers) revealed highly significant correlations with high r^2 values ($p<0.001$ in all cases; Supplemental Fig. 1). Together, these observations indicate that the global changes in AM gene expression due to smoking in healthy individuals are modest relative to the overall biological variability among healthy human individuals.

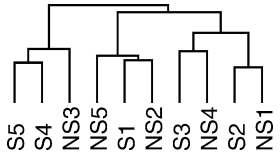

Fig. 1 The relationship of global gene expression in alveolar macrophages among different individuals. Normalized expression levels for each of the 4,199 expressed genes were analyzed using the Cluster program by hierarchical clustering with the complete linkage clustering method [16]. The clustered data was then exported to TreeView for visualization [16]. Nonsmokers are indicated as *NS*X, and smokers as *S*X, where X is the number assigned to the individual. This assessment did not segregate smokers from nonsmokers, indicating that the effects of smoking on global patterns of gene expression in alveolar macrophages are modest compared to the overall biological variability among individuals

Smoking alters the expression levels of specific genes in the alveolar macrophages of healthy smokers compared to nonsmokers

Assessment of the effect of smoking on the expression level of the 4,199 genes expressed in AM demonstrated that there were 40 up-regulated genes with $p<0.05$, and expression levels at least two-fold greater in smokers than in nonsmokers, and 35 down-regulated genes with $p<0.05$, and expression levels at least two-fold lower in smokers than in nonsmokers. Database searches (e.g., PubMed, LocusLink, GenBank, GeneCards, and OMIM) were used to determine which genes had been identified as having altered expression levels in previous studies of cigarette smoking or COPD, and to assign genes to established functional categories.

Of the 40 genes that were up-regulated in the alveolar macrophage smokers, 10 have been previously associated with the lungs in smoking and/or COPD (Table 1). Of these, six have been previously linked specifically to AM in association with smoking/COPD, and four have been previously linked to lung tissue, lung cells, or lung lavage fluid, but not specifically AM in association with smoking/COPD (Table 1).

Among the six genes that have been identified previously as being up-regulated specifically in AM in association with smoking/COPD, four were in the immune/inflammatory response category, and two genes were categorized as proteases/antiproteases, including matrix metalloproteinase 12 (macrophage elastase) and α2-macroglobulin.

Of the four genes previously identified as being up-regulated in the lungs in smoking/COPD, but not specifically linked to AM, three were in the antioxidant-related category and one was in the transcription regulation category.

Of the 35 genes that were down-regulated in smokers, only three have been previously associated with smoking and/or COPD, and none of these have been specifically linked to AM. Included among these three genes was one gene related to the immune response category, one encoding an extracellular matrix protein, and the gene encoding the 70 kD heat shock protein 2 (Table 1).

Thus, most of the genes that are differentially expressed in the AM of smokers represent novel findings. Thirty-four of the 40 genes (85%) that were observed to be up-regulated in the AM of smokers have not previously been specifically linked to AM in association with smoking/COPD (including four previously linked to smoking/COPD, but not AM per se; Tables 1 and 2). Of the 35 genes observed to be down-regulated in the AM of smokers, 32 (89%) have never been linked to the lungs in any fashion in association with smoking/COPD (including three previously linked to smoking/COPD but not to AM). Among 69 of the 75 (92%) genes up- or down-regulated in the AM of the smokers that have not been previously linked to AM or to the lungs in smoking/COPD, 10 are involved in immune responses, 6 are associated with adhesion or extracellular matrix, 9 are in the category of

Table 1 Genes previously noted to be associated with smoking and/or COPD that were observed to be up- and down-regulated in the alveolar macrophages of smokers compared to nonsmokers

Category	Gene ID[1]	Description	Gene symbol	Smokers/nonsmokers		p value[2]	Reference
				Fold up-regulated	Fold down-regulated		
Immune response and inflammation	HG4069-HT4339	Chemokine (C-C motif) ligand 2 (MCP-1)[3]	CCL2[4]	4.9		0.002	[13, 18–20]
	HG1723-HT1729	Macrophage scavenger receptor (scavenger receptor type A)	MSR1[4]	3.6		0.025	[13]
	M98399	CD36 antigen	CD36[4]	2.4		0.043	[13]
	HG1155-HT4822	Colony-stimulating factor 1	CSF1[4]	2.3		0.041	[21]
	X99133	Lipocalin 2	LCN2[5]		7.2	0.042	[26–28]
Proteolysis/ antiproteolysis	L23808	Matrix metalloproteinase 12 (macrophage elastase)	MMP12[4]	3.5		0.038	[6, 7, 22]
	M11313	α2-macroglobulin	A2M[4]	3.2		0.036	[23]
Antioxidant-related	X15722	Glutathione reductase	GSR[5]	3.2		0.047	[45]
	X55448	Glucose-6-phosphate dehydrogenase	G6PD[5]	2.1		0.025	[25]
	U30255	Phosphogluconate dehydrogenase	PGD[5]	2.1		0.011	[25]
Regulation of transcription	M84820	β-retinoid X receptor	RXRB[5]	2.2		0.037	[38–40]
Extracellular matrix	Z26653	α2-laminin	LAMA2[5]		3.9	0.031	[46]
Other	L26336	70 kDa heat shock protein 2	HSPA2[5]		3.8	0.045	[47, 48]

[1]The Gene ID numbers starting with a letter followed by five numbers are GenBank accession numbers, those starting with the letters HG are The Institute for Genomic Research (TIGR) identifiers
[2]Expression levels in alveolar macrophages of five nonsmokers and five smokers were compared using two-tailed Student's t test assuming variances
[3]*MCP-1* monocyte chemoattractant protein 1
[4]Observed in prior studies in alveolar macrophages
[5]Observed in prior studies of lung cells, tissue, or fluids other than alveolar macrophages

signal transduction, 3 encode proteases or antiproteases, 3 are involved in lysosomal function, 6 are involved in the regulation of transcription, 4 belong to the category of antioxidant-related enzymes, and 28 are involved in other functions or have not been well-characterized previously (Tables 1 and 2).

Confirmation of microarray results using TaqMan real-time reverse transcriptase polymerase chain reaction

To confirm the results obtained with microarray methodology, the expression levels of three of the differentially expressed genes that represent novel findings with respect to smoking were assessed by an independent method of RNA quantification, TaqMan real-time RT-PCR, using the same RNA samples that were used for microarray analysis. Comparisons of relative expression levels for the two genes that were up-regulated in smokers (osteopontin and ADAM10) and one gene that was down-regulated in smokers [chemokine (C-X-C motif) ligand 6] confirmed the validity of the microarray results for these genes (Fig. 2). These genes were selected for confirmation because they represented novel findings, but also because of their potential role in alveolar macrophage function (see Discussion). A two-way ANOVA confirmed that for each of the three genes there was a statistically significant effect of smoking ($p<0.01$ in all cases), but not methodology ($p>0.7$ in all cases), on expression levels. There was no significant effect of the interaction between smoking and methodology (microarray vs TaqMan) on expression levels of any of these three genes ($p>0.4$ in all cases).

Inter-individual variability in gene expression levels

Analysis of the pattern of expression by hierarchical clustering analysis of the 75 genes that were differentially expressed in smokers compared to nonsmokers suggested interindividual variability in expression levels within the two groups (Supplemental Fig. 2). For example, the serine protease member five of clade B of the serpin family (serpin B5; also known as maspin), and the gene encoding the ras-related gene associated with diabetes (RRAD) were markedly down-regulated in four of the five smokers, but expression levels of these genes in smoker one (S1) was similar to those found in nonsmokers. While S1 clustered

Table 2 Genes up- and down-regulated in the alveolar macrophages of smokers compared to nonsmokers that have not been previously associated with smoking and/or COPD

Category	Gene ID[1]	Description	Gene symbol	Smokers/nonsmokers		p value[2]
				Fold up-regulated	Fold down-regulated	
Immune response and inflammation	U20758	Osteopontin (secreted phosphoprotein 1)	SPP1	5.5		0.001
	U95626	Chemokine (C-C motif) receptor 5	CCR5	4.8		0.006
	D83920	Ficolin 1	FCN1	2.7		0.032
	U18259	Major histocompatibility complex class II transactivator	MHC2TA	2.1		0.047
	U83303	Chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	CXCL6		5.2	0.001
	L05512	Histatin 1	HTN1		3.0	0.022
	M30818	Myxovirus resistance 2	MX2		2.3	0.033
	X57351	Interferon induced transmembrane protein 3	IFITM3		2.0	0.025
	M14058	Complement component 1, r subcomponent	C1R		2.0	0.028
Adhesion and extracellular matrix	L25851	αE-integrin	ITGAE	2.4		0.008
	X15882	α2-collagen, type VI	COL6A2	2.3		0.015
	L38608	Activated leukocyte cell adhesion molecule	ALCAM	2.1		0.010
	M33308	Vinculin	VCL	2.0		0.020
	X69819	Intercellular adhesion molecule 3	ICAM3		2.4	0.045
Signal transduction	Y09561	Purinergic receptor P2X (ligand-gated ion channel 7)	P2RX7	3.4		0.036
	HG1996–HT2044	Rap2	RAP2A	2.8		0.004
	D50640	cyclic guanosine monophosphate-inhibited phosphodiesterase 3B	PDE3B	2.7		0.005
	L07597	90 kDa polypeptide 1 of ribosomal protein S6 kinase	RPS6KA1	2.1		0.003
	HG511–HT511	mitogen-activated protein kinase associated protein 1	MAPKAP1	2.1		0.004
	L24564	Ras-related associated with diabetes	RRAD		22.3	0.043
	X06182	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	KIT		3.3	0.026
	D86969	PHD finger protein 16	PHF16		2.6	0.040
	M64572	Nonreceptor type 3 protein tyrosine phosphatase	PTPN3		2.4	0.028
Proteolysis/ antiproteolysis	Z48579	Disintegrin and metalloproteinase domain 10	ADAM10	2.3		0.029
	M21188	Insulin-degrading enzyme	IDE	2.1		0.018
	U04313	Serine (or cysteine) proteinase inhibitor 5	SERPINB5		18.3	0.016
Lysosomal function	Z31690	Lipase A	LIPA	3.4		0.015
	J03263	Lysosomal-associated membrane protein 1	LAMP1	3.1		0.036
	M29877	α-L fucosidase 1	FUCA1	2.4		0.046
Regulation of transcription	L01042	TATA element modulatory factor 1	TMF1	2.4		0.048
	X59841	Pre-B cell leukemia transcription factor 3	PBX3	2.0		0.012
	L19314	Hairy and enhancer of split 1	HES1		10.0	0.027
	U44754	43 kDa polypeptide 1 of the small nuclear RNA activating complex	SNAPC1		3.0	0.049
	U09413	Zinc finger protein 135	ZNF135		2.4	0.040
Antioxidant-related	U62389	Soluble [nicotinamide adenine denucleotide phosphate (oxidized form)] isocitrate dehydrogenase 1	IDH1	2.5		0.025

Table 2 (continued)

Category	Gene ID[1]	Description	Gene symbol	Smokers/nonsmokers		p value[2]
				Fold up-regulated	Fold down-regulated	
Other	L27943	Cytidine deaminase	CDA	4.4		0.001
	L76465	Hydroxyprostaglandin dehydrogenase 15-(nicotinamide adenine dinucleotide)	HPGD	3.3		0.038
	L10381	Ribonuclease L	RNASEL	3.1		0.021
	L31529	β1-syntrophin	SNTB1	2.5		0.045
	U57623	Fatty acid binding protein 3	FABP3	2.9		0.019
	U66036	Sulfotransferase family, 1C1	SULT1C1	2.4		0.020
	U18009	Vesicle amine transport protein 1 homolog (*Torpedo californica*)	VAT1	2.4		0.027
	Z67743	Chloride channel 7	CLCN7	2.2		0.019
	M74525	Ubiquitin-conjugating enzyme E2B (RAD6 homolog)	UBE2B	2.2		0.050
	M18533	Dystrophin	DMD		8.5	0.024
	Z19574	Keratin 17	KRT17		8.4	0.037
	M13955	Keratin 7	KRT7		4.6	0.017
	HG3945–HT4215	Phospholipid transfer protein	PLTP		5.2	0.040
	X01630	Argininosuccinate synthetase	ASS		5.1	0.028
	AB002366		KIAA0368		5.0	0.026
	S58544	Sperm associated antigen 1	SPAG1		4.5	0.024
	U68385	Myeloid ecotropic viral integration site 1 homolog 4 (mouse)	MEIS4		4.4	0.030
	M19309	Slow skeletal troponin T1	TNNT1		4.4	0.044
	M64936	Retinoic acid-inducible endogenous retroviral DNA	HUMRIRT		3.8	0.026
	U02081	Neuroepithelial cell transforming gene 1	NET1		3.8	0.028
	U23070	Bone morphogenetic protein and activin membrane-bound inhibitor homolog (*Xenopus laevis*)	BAMBI		2.8	0.001
	M81883	Glutamate decarboxylase 1	GAD1		2.6	0.013
	U90716	Coxsackie virus and adenovirus receptor	CXADR		2.6	0.043
	L38933	GT198, complete open reading frame	HUMGT198A		2.4	0.017
	U60975	Sortilin-related receptor L (DLR class, A repeats-containing)	SORL1		2.4	0.000
	D16350	SA hypertension-associated homolog (rat)	SAH		2.2	0.047
	U21936	Solute carrier family 15 (oligopeptide transporter), member 1	SLC15A1		2.2	0.002

[1]The Gene ID numbers starting with a letter followed by five numbers are GenBank accession numbers, those starting with the letters HG are TIGR identifiers

[2]Expression levels in alveolar macrophages of five nonsmokers and five smokers were compared using two-tailed Student's t test assuming equal variances

with the other smokers, this individual's pattern of expression was also the most different from the other smokers, as attested by its assignment to its own branch by the clustering program (Supplemental Fig. 2). These data suggest that the levels of up- and down-regulation of gene expression in the AM of healthy smokers are variable among individual smokers, with subgroups of individuals showing similar patterns for specific groups of genes.

Discussion

The present study focuses on the identification of genes modulated by smoking in the AM of phenotypically normal smokers, with the objective of finding genes not previously associated with COPD or cigarette smoking that may play a role in the early steps of pathogenesis of this disorder. Given that COPD is characterized by chronic

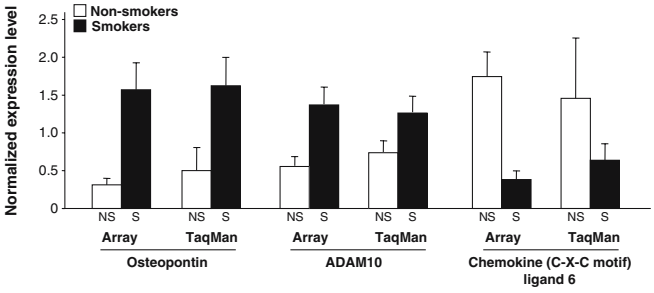

Fig. 2 Examples of the confirmation of microarray results with TaqMan real-time RT-PCR. For five nonsmokers and five smokers, expression levels in alveolar macrophages were quantified using microarrays. For all five nonsmokers and four of the five smokers, expression levels of three genes were also measured using TaqMan real-time RT-PCR on the same RNA samples used for the microarrays. To allow direct comparisons of values obtained using the two independent methods, expression levels were normalized separately for the microarray and TaqMan analysis by dividing individual values by the average expression level of all nonsmokers (NS) and smokers (S) for that method. A two-way ANOVA with smoking status (*smokers* vs *nonsmokers*) and method (*microarray* vs *TaqMan*) as independent factors confirmed that expression levels of these three genes were significantly affected by smoking status ($p<0.01$, all cases), and that method was not a significant factor ($p>0.7$, all cases). Relative expression levels are shown for osteopontin and ADAM10, genes with greater expression levels in smokers compared to nonsmokers, and chemokine (C-X-C motif) ligand 6 (CXCL6), a gene with greater expression levels in nonsmokers compared to smokers

inflammation and elevated numbers of AM in the lungs, our approach was to use microarray technology to carry out an unbiased assessment of AM gene expression in normal smokers compared to nonsmokers. The rationale in using phenotypically normal smokers with an average smoking history of 20 pack-years was two-fold: (1) using phenotypically normal smokers obviates the interpretation of the results from being complicated by secondary processes associated with responses to COPD per se, and (2) 20 pack-years is on the cusp of epidemiologic data associating the extent of smoking with the increased incidence of COPD [4–7]. The data demonstrates that smoking does not cause extensive changes in the global gene expression pattern of AM, but affects the expression of a number of specific genes. Either the gene product or the gene itself had been previously shown to be altered by smoking and/or COPD in the lungs (lung tissue, epithelium, sputum, or blood) for only 13 of the 75 genes shown to be modulated by smoking in our analysis, and of those 13, only 6 have been shown to be specifically related to AM (Table 1). The 69 novel genes observed to be up- or down-regulated in association with smoking/COPD in AM belong to defined functional categories that are important for macrophage function, including immune/inflammatory responses, adhesion and the extracellular matrix, proteolysis and antiproteolysis, signal transduction, transcription factors, and antioxidant-related function.

Genes previously noted to be associated with smoking and/or chronic obstructive pulmonary disease

The use of microarray technology allows the unbiased assessment of thousands of genes simultaneously using small amounts of biological materials. However, as with any high throughput technology, the number of false positives is a concern. One measure of the validity of results obtained from microarray analysis is consistency with previous studies that employed other methodologies to examine changes in gene expression [17]. In this context, our analysis verified the up-regulation in AM of six genes that have been identified in previous studies of AM as linked to smoking and/or COPD [6, 7, 13, 18–23].

Among the immune response and inflammatory genes noted in the present study to be up-regulated in AM in association with smoking were chemokine (CC motif) ligand-2 [CCL2, also known as monocyte chemoattractant protein 1 (MCP-1)] and colony stimulating factor 1 [CSF1, also known as macrophage colony stimulating factor (M-CSF)]. MCP-1 mRNA levels have been shown to be elevated in lung tissue sections of smokers/ex-smokers with COPD [19]. Increased MCP-1 levels have also been observed in the sputum and in BAL fluid of individuals with COPD [18, 20]. M-CSF has been noted to be up-regulated in AM isolated from healthy smokers [21]. The observation in the present study of up-regulation in the AM of smokers of the scavenger receptor type A and cluster of differentiation 36 (CD36) antigen is consistent with previously published observations linking these genes/gene products with smoking, including data showing elevated levels of CD36 in the BAL cells of smokers compared to nonsmokers.

Our observation of 3.5-fold up-regulation of the macrophage-produced matrix metalloprotease 12 (MMP12) (macrophage elastase) in cigarette smokers is in agreement with published reports linking MMP12 to the pathogenesis of emphysema [6]. MMP12 has been found to be expressed in the AM of normal cigarette smokers and of patients with emphysema [22], and single nucleotide polymorphisms in the MMP12 gene have been linked to susceptibility to smoking-induced emphysema [7]. MMP12-knockout mice exposed to cigarette smoke do not develop emphysema [6]. The data also shows up-regulation of α2-macroglobulin (α2M) expression in the AM of smokers. α2M is a protease inhibitor and cytokine transporter that inhibits many types of proteases, including collagenases and elastases [24]. The levels of α2M in culture supernatants from the AM of smokers is approximately five-fold greater than that in the AM of nonsmokers [23], and α2M is increased in the plasma of smokers compared to nonsmokers.

Interestingly, we also found up-regulation of three antioxidant-related genes (glutathione reductase, glucose-6-phosphate dehydrogenase, and phosphogluconate dehydrogenase) in the AM of smokers, in agreement with the concept that up-regulation of genes coding for antioxidant

enzymes is a likely protective mechanism of the lungs against the oxidative stress of cigarette smoke [3]. We have previously observed the up-regulation of 16 antioxidant-related genes in the airway epithelium of phenotypically normal smokers compared to nonsmokers [25], including the three antioxidant-related genes noted to be up-regulated in the AM of smokers in the present study.

Novel genes not previously associated with smoking and/or chronic obstructive pulmonary disease

In addition to identifying six smoking-modulated genes which had been previously identified in AM in association with smoking and/or COPD, microarray analysis identified 62 genes not previously linked to smoking or COPD at all, plus seven genes previously associated with smoking and/or COPD but not directly linked with AM. Because a detailed discussion of the potential role of each of these 69 genes in smoking-induced COPD is beyond the scope of this study, their possible relevance to COPD will be discussed in the context of the relevant functional categories.

Immune response and inflammation Nine genes in this category have not been previously associated with smoking and/or COPD in any fashion, and one was previously observed in regard to COPD but not in AM (lipocalin 2, a marker of neutrophil activation) [26–28]. Of the novel genes identified in our study, osteopontin had the highest differential expression in smokers vs nonsmokers (>5-fold), an observation that was confirmed by TaqMan real-time RT-PCR. Also known as secreted phosphoprotein 1, osteopontin is an arginine-glycine-aspartic acid motif-containing protein that is produced by different cell types, including activated macrophages, T cells, and osteoclasts [29–31]. This multifunctional protein has a role in chemotaxis, cell adhesion and proliferation of macrophages, smooth muscle cells, and epithelial cells [29–31]. Activated alveolar macrophage-produced osteopontin has been implicated as a fibrogenic cytokine in a bleomycin-induced mouse model of lung fibrosis [32, 33]. In contrast to the marked up-regulation of osteopontin, chemokine CXC ligand 6 (CXCL6; also known as granulocyte chemotactic protein 2) was significantly down-regulated in the AM of smokers as compared to nonsmokers. CXCL6 is a chemotactic factor for granulocytes [34]; it is possible that the down-regulation of CXCL6 is part of the mechanism of defense against inflammation in smokers.

Adhesion/extracellular matrix Our analysis identified five genes in the category of adhesion and extracellular matrix (four up-regulated: α_E-integrin, α2-collagen type 6, vinculin, and activated leukocyte cell adhesion molecule; one down-regulated: intracellular adhesion molecule 3) not previously associated with smoking/COPD, suggesting that the extent of modulation of expression levels in adhesion/extracellular matrix genes by cigarette smoke may be more widespread than previously thought.

Proteases and antiproteases The up-regulation of proteolytic enzymes by AM in response to smoking has been postulated as pivotal in the pathogenesis of COPD, in particular in the alveolar destruction that characterizes emphysema [10, 35, 36]. In addition to confirming previous observations of up-regulation of MMP12, a protease previously associated with smoking/COPD [6, 7], we also observed two proteases to be up-regulated in the AM of the smokers that have not been previously noted, including ADAM10, a member of the disintegrin and metalloprotease domain family of proteases and insulin-degrading enzymes. This observation was confirmed by TaqMan RT-PCR. In another study derived from this observation, we have shown that overexpression of ADAM10 in the lungs of mice is associated with the development of emphysema [37]. Interestingly, we also found that serpin B5, a serine protease inhibitor (serpin), member 5 of clade B of the serpin family (also known as maspin), was markedly down-regulated (8.5-fold) in the AM of smokers. This serpin has been postulated to act as a tumor suppressor, by inhibiting cell motility, adhesion, and metastasis.

Signal transduction and regulation of transcription We found that smoking modulates the expression in AM of nine genes belonging to the general category of signal transducers and six genes encoding proteins involved in transcriptional regulation. With the exception of the transcription factor β-retinoid X receptor, previously linked to smoking in lung cancer precursor lesions [38] and in the bronchial epithelium of heavy smokers [39, 40], these genes have not been previously related to smoking and/or COPD. The observations that smoking alters the expression levels of genes coding for signal transduction or transcription factors is not surprising, because cigarette smoke, like other environmental stimuli, has been reported to affect signal transduction pathways and transcription factors in the lungs, including the mitogen activated protein kinase (MAPK) pathway [14, 41–43]. Among the transcription factors, we observed down-regulation of the hairy and enhancer of split 1 (HES-1). Enhanced expression of the HES-1 gene in the early stages of macrophage development and differentiation leads to inhibition of the growth of macrophage progenitors [44]; and thus, the decreased expression of HES-1 in smokers may be related to the increased numbers of AM in the lungs of smokers [8, 9].

Overall implications

The present study provides further evidence that smoking specifically modulates alveolar macrophage gene expression in phenotypically normal smokers, before there is any evidence of lung disease. COPD is a complex disease, and there are likely multiple mediators involved in the evolution of the COPD phenotype in individuals who smoke. Microarray analysis is likely to aid in the identification of genes involved in the pathological processes leading to

smoking-induced lung diseases mediated by changes in alveolar macrophage physiology. Microarray analysis is also likely to provide novel candidate genes for susceptibility to smoking-related COPD and other lung diseases, and novel targets for therapeutic intervention. As the information evolves as to the hierarchy of mediators that dominate in the pathogenic process, microarray analysis of AM may someday be useful in predicting the development of COPD in the asymptomatic smoker.

Acknowledgements We thank N Mohamed for help in preparing this manuscript. These studies were supported, in part, by R01 HL074326-01; M01RR00047; and the Will Rogers Memorial Fund.

References

1. American Thoracic Society (1996) Cigarette smoking and health. Am J Respir Crit Care Med 153:861–865
2. Pauwels RA, Rabe KF (2004) Burden and clinical features of chronic obstructive pulmonary disease (COPD). Lancet 364:613–620
3. MacNee W (2000) Oxidants/antioxidants and COPD. Chest 117:303S–317S
4. Sethi JM, Rochester CL (2000) Smoking and chronic obstructive pulmonary disease. Clin Chest Med 21:67–86, viii
5. American Thoracic Society (1995) Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 152:S77–S121
6. Hautamaki RD, Kobayashi DK, Senior RM, Shapiro SD (1997) Requirement for macrophage elastase for cigarette smoke-induced emphysema in mice. Science 277:2002–2004
7. Joos L, He JQ, Shepherdson MB, Connett JE, Anthonisen NR, Pare PD, Sandford AJ (2002) The role of matrix metalloproteinase polymorphisms in the rate of decline in lung function. Hum Mol Genet 11:569–576
8. Bezdicek P, Crystal RG (1997) Pulmonary macrophages. In: Crystal RG, West JB, Weibel ER, Barnes PJ (eds) The lung: scientific foundations, 2nd edn. Lippincott-Raven Publishers, Philadelphia, pp 859–875
9. Shapiro SD (1999) The macrophage in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 160:S29–S32
10. Barnes PJ (2003) New concepts in chronic obstructive pulmonary disease. Annu Rev Med 54:113–129
11. Reilly JJ, Chapman HA (1988) Association between alveolar macrophage plasminogen activator activity and indices of lung function in young cigarette smokers. Am Rev Respir Dis 138:1422–1428
12. Wright JL, Hobson JE, Wiggs B, Pare PD, Hogg JC (1988) Airway inflammation and peribronchiolar attachments in the lungs of nonsmokers, current and ex-smokers. Lung 166:277–286
13. Kirkham PA, Spooner G, Ffoulkes-Jones C, Calvez R (2003) Cigarette smoke triggers macrophage adhesion and activation: role of lipid peroxidation products and scavenger receptor. Free Radic Biol Med 35:697–710
14. Koch A, Giembycz M, Stirling RG, Lim S, Adcock I, Wassermann K, Erdmann E, Chung KF (2004) Effect of smoking on MAP kinase-induced modulation of IL-8 in human alveolar macrophages. Eur Respir J 23:805–812
15. Russi TJ, Crystal RG (1997) Use of bronchoalveolar lavage and airway brushing to investigate the human lung. In: Crystal RG, West JB, Weibel ER, Barnes PJ (eds) The lung: scientific foundations, 2nd edn. Lippincott-Raven Publishers, Philadelphia, pp 371–382
16. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868
17. Chuaqui RF, Bonner RF, Best CJ, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan WM, Knezevic V, Emmert-Buck MR (2002) Post-analysis follow-up and validation of microarray experiments. Nat Genet (Suppl) 32:509–514
18. Capelli A, Di Stefano A, Gnemmi I, Balbo P, Cerutti CG, Balbi B, Lusuardi M, Donner CF (1999) Increased MCP-1 and MIP-1beta in bronchoalveolar lavage fluid of chronic bronchitis. Eur Respir J 14:160–165
19. de Boer WI, Sont JK, van Schadewijk A, Stolk J, van Krieken JH, Hiemstra PS (2000) Monocyte chemoattractant protein 1, interleukin 8, and chronic airways inflammation in COPD. J Pathol 190:619–626
20. Traves SL, Culpitt SV, Russell RE, Barnes PJ, Donnelly LE (2002) Increased levels of the chemokines GROalpha and MCP-1 in sputum samples from patients with COPD. Thorax 57:590–595
21. Rose RM, Kobzik L, Filderman AE, Vermeulen MW, Dushay K, Donahue RE (1992) Characterization of colony stimulating factor activity in the human respiratory tract. Comparison of healthy smokers and nonsmokers. Am Rev Respir Dis 145:394–399
22. Finlay GA, O'Driscoll LR, Russell KJ, D'Arcy EM, Masterson JB, Fitz Gerald MX, O'Connor CM (1997) Matrix metalloproteinase expression and production by alveolar macrophages in emphysema. Am J Respir Crit Care Med 156:240–247
23. Fujita J, Skold CM, Daughton DM, Ertl RF, Takahara J, Rennard SI (1999) Modulation of elastase binding to elastin by human alveolar macrophage-derived lipids. Am J Respir Crit Care Med 160:802–807
24. Chu CT, Howard GC, Misra UK, Pizzo SV (1994) Alpha 2-macroglobulin: a sensor for proteolysis. Ann N Y Acad Sci 737:291–307
25. Hackett NR, Heguy A, Harvey BG, O'Connor TP, Luettich K, Flieder DB, Kaplan R, Crystal RG (2003) Variability of antioxidant-related gene expression in the airway epithelium of cigarette smokers. Am J Respir Cell Mol Biol 29:331–343
26. Betsuyaku T, Nishimura M, Takeyabu K, Tanino M, Venge P, Xu S, Kawakami Y (1999) Neutrophil granule proteins in bronchoalveolar lavage fluid from subjects with subclinical emphysema. Am J Respir Crit Care Med 159:1985–1991
27. Ekberg-Jansson A, Andersson B, Bake B, Boijsen M, Enanden I, Rosengren A, Skoogh BE, Tylen U, Venge P, Lofdahl CG (2001) Neutrophil-associated activation markers in healthy smokers relates to a fall in DL(CO) and to emphysematous changes on high resolution CT. Respir Med 95:363–373
28. Keatings VM, Barnes PJ (1997) Granulocyte activation markers in induced sputum: comparison between chronic obstructive pulmonary disease, asthma, and normal subjects. Am J Respir Crit Care Med 155:449–453
29. O'Regan A (2003) The role of osteopontin in lung disease. Cytokine Growth Factor Rev 14:479–488
30. Gravallese EM (2003) Osteopontin: a bridge between bone and the immune system. J Clin Invest 112:147–149
31. Mazzali M, Kipari T, Ophascharoensuk V, Wesson JA, Johnson R, Hughes J (2002) Osteopontin—a molecule for all seasons. QJM 95:3–13
32. Takahashi F, Takahashi K, Okazaki T, Maeda K, Ienaga H, Maeda M, Kon S, Uede T, Fukuchi Y (2001) Role of osteopontin in the pathogenesis of bleomycin-induced pulmonary fibrosis. Am J Respir Cell Mol Biol 24:264–271
33. Berman JS, Serlin D, Li X, Whitley G, Hayes J, Rishikof DC, Ricupero DA, Liaw L, Goetschkes M, O'Regan AW (2004) Altered bleomycin-induced lung fibrosis in osteopontin-deficient mice. Am J Physiol Lung Cell Mol Physiol 286:L1311–L1318

34. Van Damme J, Wuyts A, Froyen G, Van Coillie E, Struyf S, Billiau A, Proost P, Wang JM, Opdenakker G (1997) Granulocyte chemotactic protein-2 and related CXC chemokines: from gene regulation to receptor usage. J Leukoc Biol 62:563–569
35. Snider GL, Ciccolella DE, Morris SM, Stone PJ, Lucey EC (1991) Putative role of neutrophil elastase in the pathogenesis of emphysema. Ann N Y Acad Sci 624:45–59
36. Suki B, Lutchen KR, Ingenito EP (2003) On the progressive nature of emphysema: roles of proteases, inflammation, and mechanical forces. Am J Respir Crit Care Med 168:516–521
37. Saitoh H, Heguy A, O'Connor TP, Harvey BG, Leopold PL, Hackett NR, Cieciuch A, Crystal RG (2004) Adenovirus-mediated delivery of ADAM10, a novel candidate gene for COPD, results in emphysematous changes in the mouse lung. Mol Ther 9:S185
38. Martinet N, Alla F, Farre G, Labib T, Drouot H, Vidili R, Picard E, Gaube MP, Le Faou D, Siat J, Borelly J, Vermylen P, Bazarbachi T, Vignaud JM, Martinet Y (2000) Retinoic acid receptor and retinoid X receptor alterations in lung cancer precursor lesions. Cancer Res 60:2869–2875
39. Xu XC, Lee JS, Lee JJ, Morice RC, Liu X, Lippman SM, Hong WK, Lotan R (1999) Nuclear retinoid acid receptor beta in bronchial epithelium of smokers before and during chemoprevention. J Natl Cancer Inst 91:1317–1321
40. Soria JC, Xu X, Liu DD, Lee JJ, Kurie J, Morice RC, Khuri F, Mao L, Hong WK, Lotan R (2003) Retinoic acid receptor beta and telomerase catalytic subunit expression in bronchial epithelium of heavy smokers. J Natl Cancer Inst 95:165–168
41. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, Palma J, Brody JS (2004) Effects of cigarette smoke on the human airway epithelial cell transcriptome. Proc Natl Acad Sci USA 101:10143–10148
42. Ning W, Li CJ, Kaminski N, Feghali-Bostwick CA, Alber SM, Di YP, Otterbein SL, Song R, Hayashi S, Zhou Z, Pinsky DJ, Watkins SC, Pilewski JM, Sciurba FC, Peters DG, Hogg JC, Choi AM (2004) Comprehensive gene expression profiles reveal pathways related to the pathogenesis of chronic obstructive pulmonary disease. Proc Natl Acad Sci USA 101:14895–14900
43. Miura K, Bowman ED, Simon R, Peng AC, Robles AI, Jones RT, Katagiri T, He P, Mizukami H, Charboneau L, Kikuchi T, Liotta LA, Nakamura Y, Harris CC (2002) Laser capture microdissection and microarray expression analysis of lung adenocarcinoma reveals tobacco smoking- and prognosis-related molecular profiles. Cancer Res 62:3244–3250
44. Masuya M, Katayama N, Hoshino N, Nishikawa H, Sakano S, Araki H, Mitani H, Suzuki H, Miyashita H, Kobayashi K, Nishii K, Minami N, Shiku H (2002) The soluble Notch ligand, Jagged-1, inhibits proliferation of CD34+ macrophage progenitors. Int J Hematol 75:269–276
45. Rahman I, MacNee W (2000) Oxidative stress and regulation of glutathione in lung inflammation. Eur Respir J 16:534–554
46. Amin K, Ekberg-Jansson A, Lofdahl CG, Venge P (2003) Relationship between inflammatory cells and structural changes in the lungs of asymptomatic and never smokers: a biopsy study. Thorax 58:135–142
47. Pinot F, el Yaagoubi A, Christie P, Dinh-Xuan AT, Polla BS (1997) Induction of stress proteins by tobacco smoke in human monocytes: modulation by antioxidants. Cell Stress Chaperones 2:156–161
48. Vayssier M, Banzet N, Francois D, Bellmann K, Polla BS (1998) Tobacco smoke induces both apoptosis and necrosis in mammalian cells: differential effects of HSP70. Am J Physiol 275:L771–L779