Journal of Pathology
J Pathol 2006; **210**: 192–204
Published online 16 August 2006 in Wiley InterScience
(www.interscience.wiley.com) **DOI:** 10.1002/path.2039

WILEY InterScience® DISCOVER SOMETHING GREAT

Original Paper

Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers†

M Woenckhaus,[1]* L Klein-Hitpass,[2] U Grepmeier,[1] J Merk,[3] M Pfeifer,[4] PJ Wild,[1] M Bettstetter,[1] P Wuensch,[5] H Blaszyk,[6] A Hartmann,[1] F Hofstaedter[1] and W Dietmaier[1]

[1] *Department of Pathology, University of Regensburg, Germany*
[2] *Department of Cell Biology, University Hospital Essen, Germany*
[3] *Department of Thoracic Surgery, University of Regensburg, Germany*
[4] *Department of Internal Medicine II, University of Regensburg, Germany*
[5] *Department of Pathology, Municipal Hospitals of Nuernberg, Germany*
[6] *Department of Pathology, University of Vermont College of Medicine, Burlington, Vermont, USA*

**Correspondence to:*
Dr M Woenckhaus, Institute of Pathology, University of Regensburg, Franz-Josef-Strauss-Allee 11, D 93053 Regensburg, Germany.
E-mail: matthias.woenckhaus@klinik.uni-regensburg.de

†*The data were presented in part at the 88th Annual Meeting of the German Association of Pathology, Rostock, Germany, 2004.*

Received: 19 December 2005
Revised: 18 May 2006
Accepted: 10 June 2006

Abstract

Tobacco smoking is the leading cause of lung cancer worldwide. Gene expression in surgically resected and microdissected samples of non-small-cell lung cancers (18 squamous cell carcinomas and nine adenocarcinomas), matched normal bronchial epithelium, and peripheral lung tissue from both smokers ($n = 22$) and non-smokers ($n = 5$) was studied using the Affymetrix U133A array. A subset of 15 differentially regulated genes was validated by real-time PCR or immunohistochemistry. Hierarchical cluster analysis clearly distinguished between benign and malignant tissue and between squamous cell carcinomas and adenocarcinomas. The bronchial epithelium and adenocarcinomas could be divided into the two subgroups of smokers and non-smokers. By comparison of the gene expression profiles in the bronchial epithelium of non-smokers, smokers, and matched cancer tissues, it was possible to identify a signature of 23 differentially expressed genes, which might reflect early cigarette smoke-induced and cancer-relevant molecular lesions in the central bronchial epithelium of smokers. Ten of these genes are involved in xenobiotic metabolism and redox stress (eg *AKR1B10, AKR1C1*, and *MT1K*). One gene is a tumour suppressor gene (*HLF*); two genes act as oncogenes (*FGFR3* and *LMO3*); two genes are involved in matrix degradation (*MMP12* and *PTHLH*); three genes are related to cell differentiation (*SPRR1B, RTN1*, and *MUC7*); and five genes have not been well characterized to date. By comparison of the tobacco-exposed peripheral alveolar lung tissue of smokers with non-smokers and with adenocarcinomas from smokers, it was possible to identify a signature of 27 other differentially expressed genes. These genes are involved in the metabolism of xenobiotics (eg GPX2 and FMO3) and may represent cigarette smoke-induced, cancer-related molecular targets that may be utilized to identify smokers with increased risk for lung cancer.
Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Keywords: gene expression; cDNA array; NSCLC; benign bronchial epithelium; smoker and non-smoker

Introduction

Lung cancer is the leading cause of cancer-related deaths in industrialized nations, occurring predominantly in patients with long-term use of tobacco products [1]. Lung cancer is classified into two major clinico-pathological groups, namely small-cell lung cancers (SCLCs) and non-small-cell lung cancers (NSCLCs). Squamous cell carcinomas (SqCs) and adenocarcinomas (ACs) are the major histological subgroups of NSCLC. The majority of squamous cell carcinomas arise centrally in the main stem of lobular or segmental bronchi and originate from transformed bronchial epithelium. Adenocarcinomas are most frequently seen as peripheral nodules, which are assumed to originate from the periphery of the bronchioloalveolar part of the bronchial tree and the peripheral alveolar lung tissue [2].

Previous studies have revealed distinct gene expression profiles in various histological types of lung cancers [3], including clinically important patient subgroups [4], smokers, and non-smokers [5], and relative to patient survival [6]. Other studies have shown that the gene expression patterns of embryonic lung development share some molecular signatures with NSCLC. Various molecular alterations have been studied in premalignant lesions [7,8] and in the histologically normal epithelium of smokers [9–12]. Recently, persistent genomic dysregulation was found

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.
www.pathsoc.org.uk

in the bronchial epithelium of previous smokers years after smoking cessation [13].

Clinical technical advances in computed tomography (CT) have improved the likelihood of detecting NSCLCs at an earlier and potentially curable stage, but spiral CT as a screening tool is insufficient in the detection of *in situ* carcinomas and micro-invasive cancers. Knowledge of the underlying genetic basis of tobacco-related lung carcinogenesis suggests that genetic changes may provide new methods for early diagnosis and risk assessment. To find putative gene expression signatures associated with cigarette smoking and cancer, we used Affymetrix U133A gene expression arrays to compare NSCLC, matched benign bronchial epithelium, and alveolar lung tissue of smokers and non-smokers.

Materials and methods

Patients and tissue samples

Our study included only patients with no prior treatment, a known smoking history, informed consent, and sufficient amounts of frozen tumour and normal tissue for molecular analysis. Seven women and 20 men with NSCLC and without pretreatment (mean age 61 years, range 38–77 years) who underwent surgery at the University Hospital Regensburg (2001–2003) were studied (Table 1). The Institutional Ethics Committee (No 01/97) approved the study. Clinical information was obtained through chart review. One pathologist (MW) verified the histological diagnoses (nine Acs and 18 SqCs) and stage (pT1, $n = 9$ and pT2, $n = 18$; TNM classification, UICC, WHO). A current smoking history with a mean of 40 pack years (range 12–85 pack years) was found in 22 patients. In the group of squamous cell carcinomas, far more males (17 males and one female) were identified, while female patients were more often affected by adenocarcinomas (six females and three males). Five individuals (two men and three women) with adenocarcinomas were non-smokers. Because adenocarcinoma cases from non-smokers were specifically selected, we have an over-representation of women in our study.

As controls, we used benign bronchial epithelium and peripheral lung tissue from eight patients (mean age 52 years; range 33–67 years) who were treated only by surgery and who underwent lung resection for a diagnosis other than NSCLC [four patients with metastatic cancer (renal cell carcinoma, colorectal carcinoma, testicular carcinoma, and melanoma) and four patients with carcinoid tumours]. There were three smokers (mean smoking history 31 pack years; range 7–45 pack years) in this group.

RNA isolation and RNA pool preparation

Resected tumour specimens were evaluated by routine frozen section procedures and study samples were snap-frozen and stored at −80 °C. Frozen samples were cut (5 µm) and stained with haematoxylin and eosin. Using manual microdissection under an inverted microscope or laser microdissection (PALM), tumour cells, bronchial cells, and alveolar lung tissue were selected. Touch preparations were used to isolate bronchial epithelium [14]. Papanicolau-stained touch preparations were examined to verify the absence of contaminating cells in each case. Total RNA was isolated using the RNeasy Mini Kit (Qiagen Science, Maryland, USA) following the manufacturer's protocol. The RNA was treated with DNase I (Qiagen). The integrity and concentration of RNA were confirmed twice by spectrophotometry and the Agilent 2100 Bioanalyzer using RNA 6000 Nano Chips (Agilent Technologies GmbH, Waldbronn, Germany). Only high-quality RNA with two distinct ribosomal RNA bands (18S and 28S rRNA) was used for subsequent GeneChip hybridization.

Overall, we prepared 12 pools including 35 patients. RNA pools with similarity in the histological type and smoking history were combined to equal amounts of RNA (Table 2).

Table 1. Patients, histological diagnosis, tumour grade, and stage according to the TNM classification and smoking history

Patient ID	Diagnosis	Sex	Age at surgery (years)	T	N	G	Smoking history (pack years)
1	SqC	M	56	1	1	2	85
2	SqC	M	70	2	0	2	40
3	SqC	M	55	2	2	3	45
4	SqC	M	75	2	0	2	50
5	SqC	M	58	2	0	2	60
6	SqC	F	57	2	0	3	60
7	SqC	M	63	1	0	2	80
8	SqC	M	67	2	0	2	50
9	SqC	M	60	1	0	3	30
10	SqC	M	60	2	0	3	40
11	SqC	M	52	1	1	2	30
12	SqC	M	66	1	0	3	30
13	SqC	M	57	2	0	3	30
14	SqC	M	50	2	1	2	30
15	SqC	M	51	2	0	3	30
16	SqC	M	63	2	0	2	30
17	AC	M	68	1	2	2	30
18	AC	M	63	1	1	3	30
19	AC	M	38	1	0	3	40
20	AC	F	60	2	2	2	25
21	AC	F	39	2	0	3	30
22	AC	F	59	2	0	3	12
23	AC	M	67	2	0	2	NS
24	AC	F	66	2	0	2	NS
25	AC	F	71	2	0	2	NS
26	AC	F	66	1	0	2	NS
27	AC	M	77	2	0	2	NS
28	TuFr	M	50	—	—	—	NS
29	TuFr	F	47	—	—	—	NS
30	TuFr	F	67	—	—	—	NS
31	TuFr	F	38	—	—	—	NS
32	TuFr	M	50	—	—	—	NS
33	TuFr	M	66	—	—	—	45
34	TuFr	M	33	—	—	—	7
35	TuFr	M	66	—	—	—	40

SqC = squamous cell cancer; AC = adenocarcinoma; TuFr = tumour-free, without primary lung cancer; F = female; M = male.

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Table 2. RNA pool preparation

Histology	PY	Br RNA samples	Ca RNA samples	Lu RNA samples
SqC	60 PY (mean)	Pool 1 (ID: 1–7)	Pool 2 (ID: 1–8)	Pool 3 (ID: 2–7)
SqC	60 PY (mean)		Single probes (ID: 1–6)	
SqC	31 PY (mean)	Pool 4 (ID: 9–18)		
AC	31 PY (mean)	Pool 5 (ID: 19–21)	Pool 6 (ID: 20–22)	
AC	Non-smokers	Pool 7 (ID: 24–27)	Pool 8 (ID: 23–25)	Pool 9 (ID: 23–27)
AC	30 PY		Single probes (ID: 21)	
TuFr	30 PY (mean)	Pool 10 (ID: 33–35)		
TuFr	Non-smokers	Pool 11 (ID: 28–32)		Pool 12 (ID: 28–32)

SqC = squamous cell cancer; AC = adenocarcinoma; TuFr = tumour-free, without primary lung cancer; PY = pack years; Br RNA samples = bronchial RNA samples; Ca RNA samples = carcinoma RNA samples; Lu RNA samples = alveolar lung tissue RNA samples; ID = patient ID.

RNA amplification and oligonucleotide microarray analysis

A modified two-round amplification procedure [15] starting with 300 ng of RNA was used to obtain sufficient amounts of biotin-labelled cRNA suitable for array hybridization. Labelled cRNA was purified on RNeasy columns (Qiagen, Hilden, Germany), fragmented, and hybridized to HG-U133A microarrays (Affymetrix, Santa Clara, CA, USA) following the Affymetrix standard protocol. The arrays were washed and stained according to the manufacturer's recommendation and finally scanned in a GeneArray Scanner 2500 (Agilent, Palo Alto, CA, USA). Array images were processed to determine signals and detection calls (present, absent, marginal) for each probe set using Affymetrix Microarray Suite 5.0 software. Scaling across all probe sets of a given array to an average intensity of 1000 was performed to compensate for variations in the amount and quality of the cRNA samples and other experimental variables of non-biological origin.

Cluster analysis

For unsupervised hierarchical clustering, signals of individual probe sets were normalized to the median probe set signal of all included arrays and log-transformed. Log-transformed ratios were subjected to UPGMA clustering using correlation as similarity measure (Spotfire DecisionSite for functional genomics). Only probe sets receiving present detection calls in at least 30% of the arrays were evaluated.

Identification of differentially expressed genes

To identify differentially regulated transcripts, experimental and baseline arrays were analysed by pair-wise cross-comparison analysis using Affymetrix GCOSv1.2 software. Comparisons were carried out with GCOSv1.2, which calculates a Signal Log2 Ratio as well as the significance (change p value, Wilcoxon ranking test) of each change in gene expression based on the signal differences between perfect match and mismatch oligonucleotide probes of the respective probe sets on the two arrays compared. Probe sets showing greater than two-fold up- or down-regulation and a change value less than or equal to 0.0045 and at least one present detection call in the two single array analyses of the array pair were identified by filtering and defined as differentially expressed targets. Target probe sets that met these criteria in more than 50% of all possible cross-comparisons analysed were defined as consistently dysregulated.

Identification of statistically overrepresented GO terms within lists of target genes

To identify Gene Ontology categories which are overrepresented in target gene lists, we used the GOstat tool [16] (http://gostat.wehi.edu.au) and Fisher exact test for statistical significance. As a reference list, all non-AFFX control probe sets (22215) of the HG-U133A array were used. To correct for the multiple testing problem, the Benjamini correction option was chosen [17].

Real-time relative quantitative RT-PCR

Relative quantitative real-time PCR was used to confirm the differential expression of a selected number of genes. Reverse transcription was performed with 1 µg of total RNA for 50 min at 42 °C using a random hexamer primer dN_6 (Roche Diagnostics GmbH, Penzberg, Germany) and Superscript II reverse transcriptase (Life Technologies, Inc) [18]. Real-time RT-PCR analyses of eight genes were done using the LightCycler and Faststart DNA Master SYBR Green I kit (Roche Diagnostics GmbH, Penzberg, Germany). The grancalcin (GCA) gene was chosen as a reference due to its constant expression at moderate level in the cDNA array experiment. Each PCR result was normalized against GCA RNA. cDNA from the cell line HT26 was used as a calibrator. Quantification of target gene expression was carried out using Relative Quantification Software Version 1.0 (Roche Diagnostics) with efficiency correction. Primer sequences are given in Table 3.

Immunohistochemistry (IHC)

Immunohistochemical studies utilized an avidin–biotin peroxidase method with a diaminobenzidine (DAB)

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Table 3. Primers and annealing temperatures used in real-time RT-PCR

Gene	Primer	*T* (°C)*	Base pair
AKR1B10	Fwd: AACGTGTTGCAATCCTCTCA Rev: AGCTTGGCTAAAATGGGACA	62	143
ALDH3A1	Fwd: GCCAAGATGACCCAGCAC Rev: CAGGTCAGCAGAGGAGTGG	62	150
BLU	Fwd: AAGGTGGGCTGAGACCTACA Rev: ACTCCCTGCAGCAATACCAC	57	134
GAS2	Fwd: CTGGATGATGCAGTGAAACG Rev: TCCCACACGGACCATGAC	62	156
GCA	Fwd: GTGTGAAGCTTCGAGCATTG Rev: CCATAGTGCCCTGCAAAAA	62	102
MGP	Fwd: TGAATCACATGAAAGCATGG Rev: CTCGGATCCTCTCTTGGACTT	62	119
MSMB	Fwd: TGCACTTGCTACGAAACAGAA Rev: GTCCTTCTTCTCCACCACGA	62	129
SCG2	Fwd: AGAACGGGGAGGAATATGCT Rev: GCCCCAGAGATGAGGAAAAT	62	140
WIF1	Fwd: AAGGTTGGCATGGAAGACAC Rev: AAAACGTTTCAGATGTCGGAGT	62	174

* PCR annealing temperature.

chromatogen. Two-micrometre sections of formalin-fixed, paraffin-embedded tissue samples were incubated overnight and after antigen retrieval (microwave oven for 35 min at 250 W). Immunohistochemistry was carried out in a NEXES immunostainer (Ventana, Tucson, AZ, USA) following the manufacturer's instructions. The following primary antibodies were used: *WIF1* (1 : 50; as previously described [19]), cyclin A (1 : 10; Santa Cruz Biotechnology Inc, Heidelberg, Germany), MMP7 (1 : 20; Oncogen Research Products, Cambridge, MA, USA), β-catenin (1 : 50; Santa Cruz Biotechnology Inc, Heidelberg, Germany), P-cadherin (1 : 30; Novocastra, UK), FGFR3 (1 : 50; Santa Cruz Biotechnology Inc, Heidelberg, Germany), and monoclonal anti-human maspin antibody (1 : 1000; BD Biosciences Pharmingen™, San Diego, CA, USA). Negative controls without primary antibody were included in each experiment. The cytoplasmic and nuclear immunoreactivity was assessed semi-quantitatively, based on the percentage of positive cells. Weak staining in less than 30% was defined as 1+, intermediate in 30–70% as 2+, and strong staining in more than 70% as 3+; when scoring cyclin A staining, the percentage of positively stained cells per 100 cells was determined.

Results

Unsupervised hierarchical cluster analysis

The unsupervised hierarchical cluster analysis separated two main branches (Figure 1). The right main branch included squamous cell carcinomas (pool 2, single samples ID 1, 2, 4, 5, and 6) and adenocarcinomas of smokers (pool 6, sample ID 21) and non-smokers (pool 8) clustered separately. The non-malignant tissue within the left main branch dispersed in one branch including peripheral alveolar lung tissue (pools 3, 9, and 12) and another branch containing the bronchial epithelium (pools 1, 4, 5, 7, 10, and 11), which split into two sub-branches, one representing smokers (pools 1, 4, 5, and 10) and the other containing non-smokers (pools 7 and 11).

Differentially expressed probe sets in NSCLC

Differentially expressed probe sets in squamous cell cancers of smokers

To extract the cancer-related expression profiles, we compared squamous cell carcinomas with the bronchial epithelium from which they are assumed to originate (cross-comparison: pool 2 versus pool 1 and pool 4). Supplementary Table 1 (available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html) shows the differentially expressed genes. We found 2355 consistently dysregulated probe sets in squamous cell carcinoma; 974 of them were up-regulated and 1381 were down-regulated.

Differentially expressed probe sets in adenocarcinomas of smokers

Because the adenocarcinoma cells of smokers are thought to differentiate from various cells of the peripheral bronchioloalveolar tree or potentially even descend from the alveolar part of the lung, we compared separately the adenocarcinomas with the bronchial epithelium and the peripheral alveolar lung tissue. Comparison of adenocarcinomas with bronchi showed 1479 differentially expressed probe sets (Supplementary Table 2, available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html), including 342 that were up-regulated and 1137 that were down-regulated (comparison: pool 6 versus pool 5). By comparing adenocarcinomas with peripheral alveolar lung tissue, we identified

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

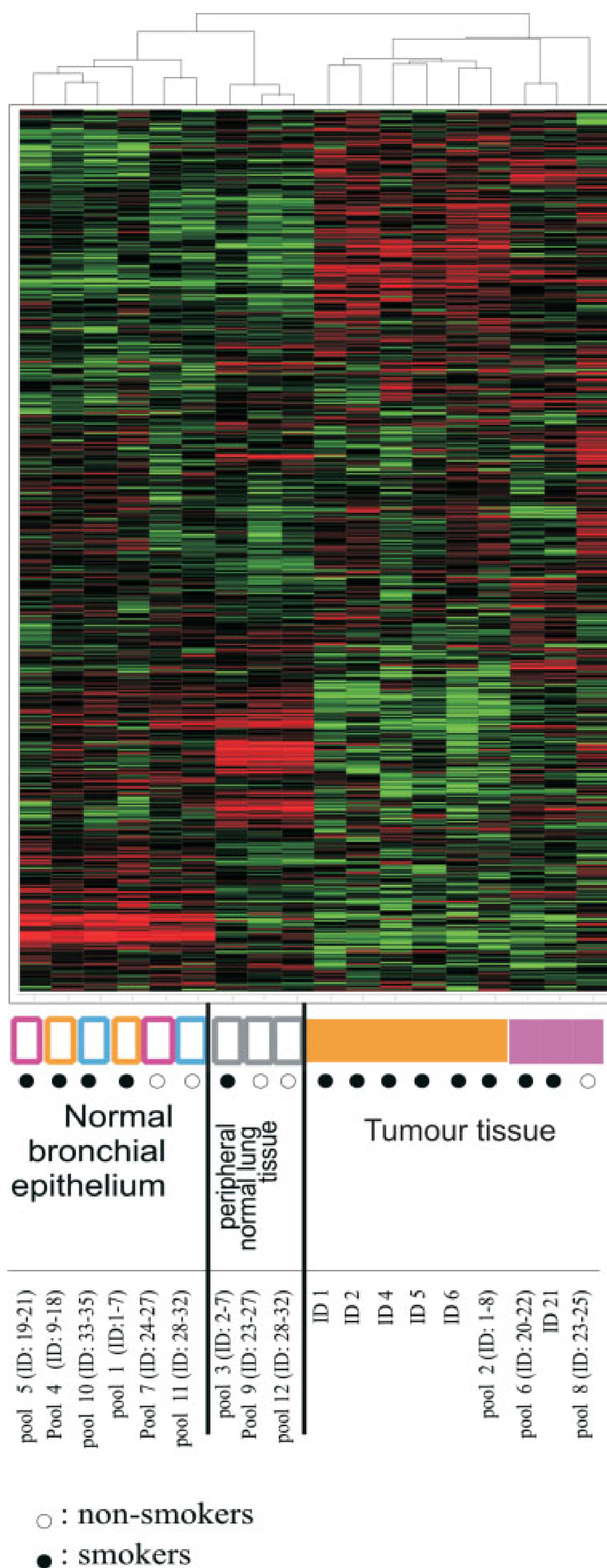

Figure 1. Two-dimensional unsupervised hierarchical clustering of 18 cDNA array analyses representing nine NSCLCs (six squamous cell carcinomas and three adenocarcinomas), six sample pools from benign bronchial epithelium, and three sample pools from peripheral lung tissue. Each row represents the expression pattern of a single gene for all samples of tissues ($n = 15\,384$ genes), and each column the expression of all genes in a tissue sample. Bars represent the type of lung cancer and the smoking history of patients from whom the tissue samples were taken. Orange, squamous cell carcinoma; pink, adenocarcinoma; blue, no primary NSCLC; grey, peripheral lung tissue; hatched, non-smoker; filled, smoker

102 differentially expressed probe sets (Supplementary Table 5, available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html), including 55 up-regulated and 47 down-regulated genes (comparison: pool 6 versus pool 3).

Differentially expressed probe sets in adenocarcinomas of smokers compared with adenocarcinomas of non-smokers

Comparing adenocarcinomas with and without a smoking history (comparison: pool 6 versus pool 8), we found 948 differentially expressed probe sets, listed in Supplementary Table 3 (available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html) (including 772 up-regulated and 176 down-regulated genes).

Differentially expressed probe sets in tobacco-exposed normal tissues

Differentially expressed probe sets in bronchial epithelium of smokers

To identify smoking-related gene expression alterations, we compared the expression profiles of the bronchial epithelium of smokers and non-smokers (cross-comparison: pools 1, 4, 5, 10 versus pools 7, 11). Dysregulated transcripts (244 genes down-regulated, 199 genes up-regulated) were seen in 443 probe sets (Supplementary Table 4, available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html). The GOstat tool software was used to identify genes with statistically overrepresented functional categories and it detected 30 differentially expressed genes assigned to cell adhesion ($p < 0.0001$; Table 4).

Differentially expressed probe sets in peripheral alveolar lung tissue of smokers

The additionally performed comparison of differentially expressed genes in compared peripheral alveolar lung tissue from smokers and non-smokers is listed in Supplementary Table 5 (available at http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html) (comparison: pool 3 versus pool 12) [102 differentially expressed probe sets (60 up-regulated, 42 down-regulated)].

Consistently differentially expressed probe sets in tobacco-exposed bronchi and peripheral alveolar lung tissue

The results presented in Table 5 are based on a double comparison. Table 5A shows dysregulated genes in smokers with squamous cell carcinoma that were in the same way up- or down-regulated in both tumour and the bronchial epithelium of smokers (comparison: single probes ID 1–6 versus pool 1 and pool 4). These genes were not dysregulated in the bronchial epithelium of non-smokers (comparison: pools 1, 4, 5, 10 versus pools 7, 11).

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Table 4. List of genes with changed expression that are significantly over- or under-represented in normal bronchus of smokers versus non-smokers (GO functional categories)

UniGene ID	Gene symbol	Gene title	Median expression value of		
			smokers (S)	Non-smokers (NS)	Ratio of S/NS
Up-regulated transcripts					
Hs.408488	MSLN	Mesothelin	253.25	46.85	5.4
Hs.436319	PARVA	Parvin, alpha	466.5	178.15	2.6
Hs.369646	MUC 4	Mucin 4, tracheobronchial	1651.8	692.7	2.4
Hs.476018	CTNNB1	Catenin, beta 1	6437.85	2836	2.3
Hs.404119	TSTA3	Tissue-specific transplantation antigen P35B	744.05	366.15	2.0
Hs.483408	PPP2CA	Protein phosphatse 2, catalytic subunit, alpha isoform	1967.1	1034.55	1.9
Down-regulated transcripts					
Hs.120949	CD36	CD36 antigen (collagen type I receptor)	465.7	873.3	0.5
Hs.120950	TGFBI	Transforming growth factor, beta-induced, 68 kD	1975.05	4047.45	0.5
Hs.120951	PECAM1	Platelet/endothelial cell adhesion molecule (CD31 antigen)	728.85	1499	0.5
Hs.41296	FLRT3	Fibronectin leucine-rich transmembrane protein 3	363.15	747.2	0.5
Hs.409034	COL15A1	Collagen, type XV, alpha 1	315.05	739.35	0.4
Hs.489646	LAMB1	Laminin, beta 1	378.35	890.45	0.4
Hs.511397	MCAM	Melanoma cell adhesion molecule	878.35	2087.95	0.4
Hs.80552	DPT	Dermatopontin	165.5	409.15	0.4
Hs.409662	COL14A1	Collagen, type XIV, alpha 1 (undulin)	317.6	788.1	0.4
Hs.440848	VWF	von Willebrand factor	1003.85	2686.75	0.4
Hs.15154	SRPX	Sushi-repeat-containing protein, X-linked	523.2	14445.2	0.4
Hs.522891	CXCL12	Chemokine (C-X-C motif) ligand 12	591.8	1717.4	0.3
Hs.196647	KIAA0527	KIAA0527 protein	45.55	135.95	0.3
Hs.131704	NRP1	Neuropilin 1	166.05	526.8	0.3
Hs.517227	JAM2	Junctional adhesion molecule 2	404.15	1284	0.3
Hs.89546	SELE	Selectin E (endothelial adhesion molecule 1)	378.2	1300.1	0.3
Hs.371147	THBS2	Thrombospondin 2	722.4	2703.15	0.3
Hs.508597	ITGBL1	Integrin, beta-like 1(with EGF-like repeat domains)	116.2	440	0.3
Hs.387794	ChGn	Chondroitin beta1,4 *N*-acetylgalactosaminyltransferase	557.55	2202.1	0.3
Hs.97199	C1QR1	Complement component 1, q subcomponent, receptor 1	939.35	3793.25	0.2
Hs.371199	SGCE	Sarcoglycan, epsilon	518.6	2762.2	0.2
Hs.523446	COL11A1	Collagen, type XI, alpha 1	111.5	826.5	0.1
Hs.489142	COL1A2	Collagen, type I, alpha 2	660.85	5653.7	0.1
Hs.2799	HAPLN1	Hyaluronan and proteoglycan link protein 1	30.75	341.7	0.1

S = smoker; NS = non-smoker.

In Table 5B, the same strategy was applied to the comparison of adenocarcinomas of smokers with equally dysregulated genes in the bronchial epithelium of smokers (comparison: pool 6 versus pool 5), but were not dysregulated in the bronchial epithelium of non-smokers (pools 1, 4, 5, 10 versus pools 7, 11).

Since adenocarcinomas could also descend from the alveolar part of the lung, a further double comparison was performed (Table 5C). Here we looked for equal dysregulation in adenocarcinomas of smokers and the peripheral alveolar lung tissue of smokers (comparison: pool 6 versus pool 3). These genes were not dysregulated in the alveolar lung tissue of non-smokers (pool 3 versus pool 12).

Verification of expression changes by relative quantitative RT-PCR

To confirm the expression levels obtained by the expression arrays with an independent method, relative quantitative real-time PCR using the LightCycler (Roche) was performed for eight different genes. These included *AKR1B10, ALDH3, BLU, GAS2MGP, MSMB, SCG2*, and *WIF1* (Figure 2). Expression of up-regulated ($n = 2$) and down-regulated ($n = 6$) genes was examined in eight individual tumour specimens and two pools (SqC, AC). Pools from bronchial epithelia and alveolar lung tissue were tested as well. The expression data generated by the Affymetrix array and relative real-time RT-PCR were highly concordant (92%), supporting the reliability of the array analyses (Figure 2).

Verification of expression changes by immunohistochemistry

Samples from similar patients that were investigated by gene expression were also used for immunohistochemistry, including WIF-1 ($n = 18$), cyclin A ($n = 22$), maspin ($n = 25$), FGFR3 ($n = 25$), P-cadherin ($n = 16$), β-catenin (n = 27), and MMP 7 (n = 27).

WIF-1 staining (Figure 3A) showed perinuclear cytoplasmic staining in the bronchi in 83% (15/18) and negative or weak staining in 72% of cancer tissues (13/18).

Cyclin A was negative in the bronchial cells and strongly up-regulated in about 30% of tumour cells (Figure 3B).

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Table 5. Consistent molecular alterations in tobacco smoke-exposed benign and malignant tissues

(A) Genes consistently dysregulated in smokers' normal bronchial epithelium (Br) and squamous cell carcinoma (SqC)

			Br (S) vs Br (NS)		SqC (S) vs Br (NS)	
UniGene ID	Gene symbol	Gene title	*n*	Regulation	*n*	Regulation
Hs. 80962	NTS	Neurotensin	7/8	Down	8/12	Down
Hs. 17109	ITM2A	Integral membrane protein-2	6/8	Down	8/12	Down
Hs. 445350	FMO3	Flavin-containing monooxygenase 3	7/8	Down	10/12	Down
Hs. 137367	ANK2	Ankyrin 2, neuronal	5/8	Down	7/12	Down
Hs. 196952	HLF	Hepatic leukaemia factor	7/8	Down	7/12	Down
Hs. 143811	APIN	Apin protein	5/8	Down	7/12	Down
Hs. 196952	HLF	Hepatic leukaemia factor	6/8	Down	8/12	Down
Hs. 4	ADH1B	Alcohol dehydrogenase IB (class I)	6/8	Down	9/12	Down
Hs. 150595	CYP26A1	Cytochrome P450, family 26, subfamily A	7/8	Down	9/12	Down
Hs. 368626	RTN1	Reticulon 1	5/8	Down	9/12	Down
Hs. 144912	FMO2	Flavin-containing monooxygenase 2	7/8	Down	9/12	Down
Hs. 103944	MUC7	Mucin 7, salivary	6/8	Down	9/12	Down
Hs. 504908	LMO3	LIM domain only 3 (rhombotin-like 2)	5/8	Down	10/12	Down
Hs. 59889	HMGCS2	Cytoplasmic HMG-CoA synthase 2	6/8	Down	10/12	Down
Hs. 28988	GLRX	Glutaredoxin	5/8	Down	10/12	Down
Hs. 116724	ALKR1B10	Aldo-keto reductase family 1, member B10	6/8	Up	7/12	Up
Hs. 89626	PTHLH	Parathyroid hormone-like hormone	7/8	Up	10/12	Up
Hs. 1076	SPRR1B	Small proline-rich protein 1B (comifin)	5/8	Up	9/12	Up
Hs. 1695	MMP12	Matrix metalloproteinase 12	5/8	Up	12/12	Up
Hs. 81134	IL1RN	Interleukin 1 receptor antagonist	5/8	Up	10/12	Up

(B) Genes consistently dysregulated in smokers' normal bronchial epithelium (Br) and adenocarcinomas (ACs)

			Br (S) vs Br (NS)		AC (S) vs Br (S)	
UniGene ID	Gene symbol	Gene Title	*n*	Regulation	*n*	Regulation
Hs. 445350	FMO3	Flavin-containing monooxygenase 3	7/8	Down	1/1	Down
Hs. 103944	MUC7	Mucin 7, salivary	6/8	Down	1/1	Down
Hs. 188518	MT1K	Metallothionein 1K	7/8	Down	1/1	Down
Hs. 13872	CYTL1	Cytokine-like 1	8/8	Down	1/1	Down
Hs. 487046	SOD2	Superoxide dismutase 2, mitochondrial	6/8	Up	1/1	Up
Hs. 498513	AKR1C1/AKR1C2	Aldo-keto reductase family 1, member C1/C2	6/8	Up	1/1	Up

(C) Genes consistently dysregulated in smokers' normal alveolar epithelium (lung) and adenocarcinomas (AC)

			Lung (S) vs lung (NS)		AC (S) vs lung (NS)	
UniGene ID	Gene symbol	Gene title	*n*	Regulation	*n*	Regulation
Hs. 2704	GPX2	Glutathione peroxidase 2 (gastrointestinal)	1/1	Up	1/1	Up
Hs. 40499	DKK1	Dickkopf homologue 1 (*Xenopus laevis*)	1/1	Up	1/1	Up
Hs. 27879	MAGEA1	Melanoma antigen family A,1	1/1	Up	1/1	Up
Hs. 271580	UPK1B	Uroplakin 1B	1/1	Up	1/1	Up
Hs. 244723	CCNE1	Cyclin E1	1/1	Up	1/1	Up
Hs. 130546	FLJ20449	Hypothetical protein FLJ20449	1/1	Up	1/1	Up
	H2 histone family	H2 histone family	1/1	Up	1/1	Up
Hs. 20833	HMOX1	Heme oxygenase (decycling) 1	1/1	Down	1/1	Down
Hs. 77348	HPGD	15-Hydroxyprostaglandin-dehydrogenase 15-(NAD)	1/1	Down	1/1	Down
Hs. 82002	EDNRB	Endothelin receptor type B	1/1	Down	1/1	Down
Hs. 78061	TCF21	Transcription factor 21	1/1	Down	1/1	Down
Hs. 646	CPA3	Carboxypeptidase A3 (mast cell)	1/1	Down	1/1	Down
Hs. 237868	IL7R	Interleukin-7 receptor	1/1	Down	1/1	Down
Hs. 94592	KL	Klotho	1/1	Down	1/1	Down
Hs. 80876	FMO3	Flavin-containing monooxygenase 3	1/1	Down	1/1	Down
M98399	CD36	CD36 (Collagen type I receptor)	1/1	Down	1/1	Down
M24317		Alcohol dehydrogenase IB (class I)	1/1	Down	1/1	Down
U58913		Chemokine (C-C motif) ligand 23	1/1	Down	1/1	Down
Hs. 202949	KIAA1102	KIAA1102 protein	1/1	Down	1/1	Down
Hs. 25956	SOSTDC1	Sclerostin domain-containing protein 1	1/1	Down	1/1	Down
Hs. 293441	IGHA1	Immunoglobin heavy constant alpha 1	1/1	Down	1/1	Down
Hs. 173233	FLJ10970	Hypothetical protein FLJ10970	1/1	Down	1/1	Down
Hs. 105468	GIMAP6	IMAP family member 6″	1/1	Down	1/1	Down
Hs. 211869	DKK2	Dickkopf homologue 2 (*Xenopus laevis*)	1/1	Down	1/1	Down
AA807056	—	Major histocompatibility complex	1/1	Down	1/1	Down
M83772	FMO3	Flavin-containing monooxygenase 3	1/1	Down	1/1	Down

S = smoker; NS = non-smoker; *n* = number.

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Nuclear maspin (SERPINB5) staining was basally localized in the bronchi in 88% (22/25) and cytoplasmic staining (2+ or 3+) was predominantly seen in 88% of squamous cell cancers (21/24). Fifty-six per cent (11/25) of SqC cases showed nuclear staining (Figure 3C).

Both FGFR3 and P-cadherin only focally stained basal bronchial cells. P-cadherin showed strong staining in only 42% (6/16) of carcinomas (Figure 3D), whereas 18% (3/16) showed heterogeneous weak positivity and 43% (7/16) were negative. MMP 7 (Figure 3E) was diffusely and heterogeneously detectable in 74% of cancers (20/27). Most squamous cell carcinomas showed intermediate (+2) or strong (+3) FGFR3 up-regulation in 78% of tumours (21/27) (Figure 3F). β-Catenin showed strong cytoplasmic staining in all bronchi and cancers of patients with a smoking history (27/27) (Figure 3G1) and weaker expression in the bronchi of non-smokers (Figure 3G2). Peripheral alveolar lung tissue was negative in all cases and immunohistochemical differences between smokers and non-smokers were not seen. All the immunohistochemical results confirmed the findings of RNA expression, except for P-cadherin, where several tumours with strongly up-regulated RNA levels were negative by immunohistochemistry.

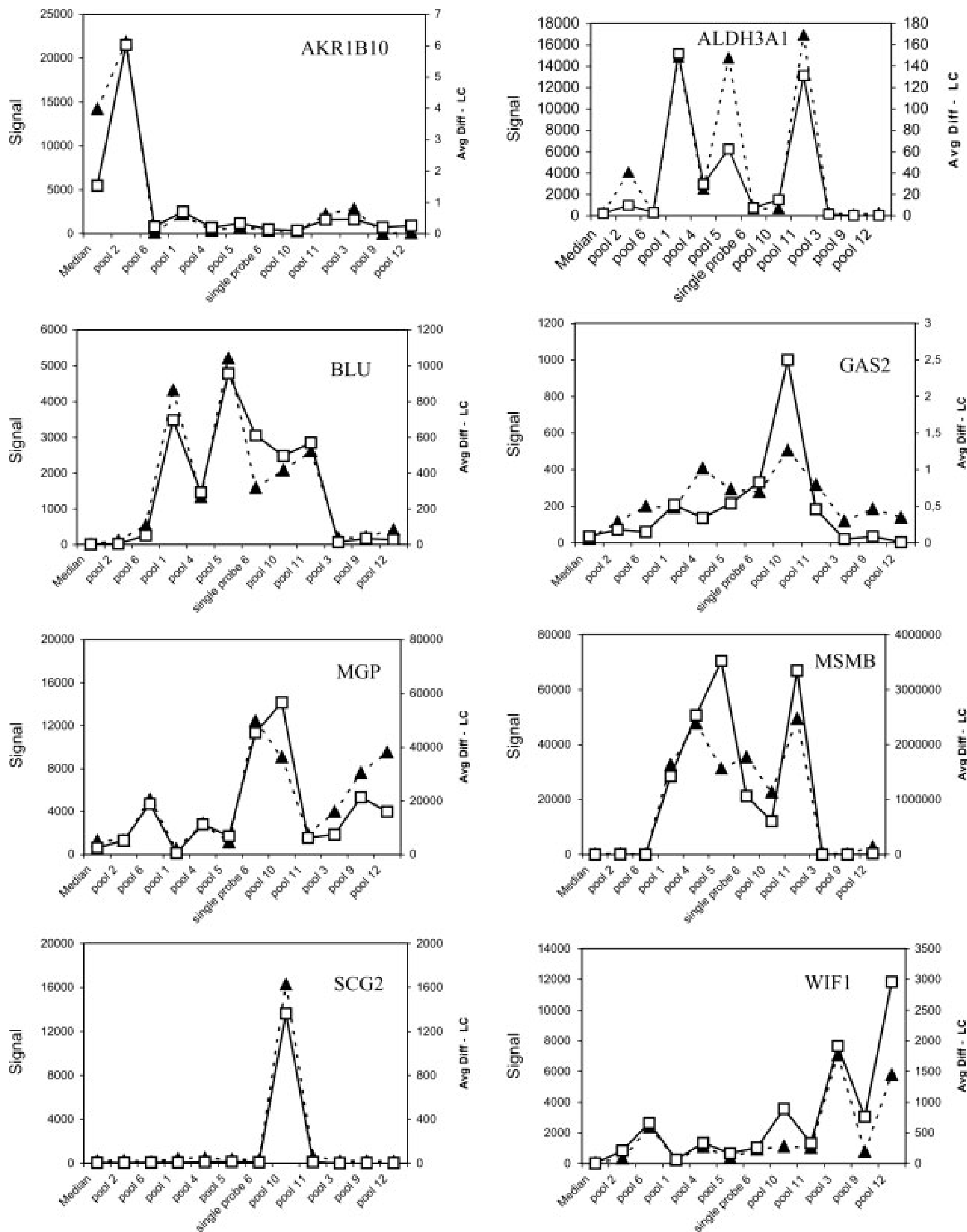

Figure 2. Verification of GeneChip analyses by real-time RT-PCR. Expression analyses of eight selected genes are shown. ▲, Expression intensity (signal) from GeneChip analyses; □, average diff-LC: average values of double determinations normalized to GCA from real-time PCR

J Pathol 2006; **210**: 192–204 DOI: 10.1002/path
Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Discussion

The present gene expression profiling study was performed to provide further insight into the molecular mechanisms of lung carcinogenesis and to identify smoking-associated and cancer-related molecular alterations of bronchial epithelium and matched NSCLC. Tissue microdissection is crucial for reliable exact molecular analysis [20,21], and manual microdissection, laser microdissection or touch preparations were used in every case to exclude contaminating non-target cells. Even though RNA isolation of the single-layer bronchial epithelium performed by laser microdissection and touch preparations generates the same amounts of RNA, the quality of the RNA samples was better using touch preparations. We used the more cell-sparing touch preparation method controlled by microscopy in bronchi because this technique yields high-quality RNA samples [14].

Epidemiological studies in patients with squamous cell carcinomas have demonstrated a predominance of males, while adenocarcinomas develop more frequently in non-smoking women [1]. Adenocarcinomas from non-smokers were specifically selected, which accounts for the overrepresentation of women in our study. Only RNA samples from cases with similar histological type and smoking status were combined. Sample pooling enabled us to investigate small bronchial samples by providing enough high-quality RNA for subsequent confirming RT-PCR. In addition, we investigated single probes because gene expression profiling of pooled samples may show the dominant expression pattern of combined samples and could repress individual alterations [22]. Hierarchical cluster analysis showed single squamous cell carcinomas within the pooled squamous differentiated branches and single adenocarcinomas beside the pools of adenocarcinomas, demonstrating the purity of microdissected pooled and single probes.

Hierarchical cluster analysis distinguished benign from malignant tissue, and squamous cell carcinomas and adenocarcinomas. The cluster arrangement was consistent with the smoking history in patients with adenocarcinomas with and without a smoking history. This matches the results of Powell *et al*

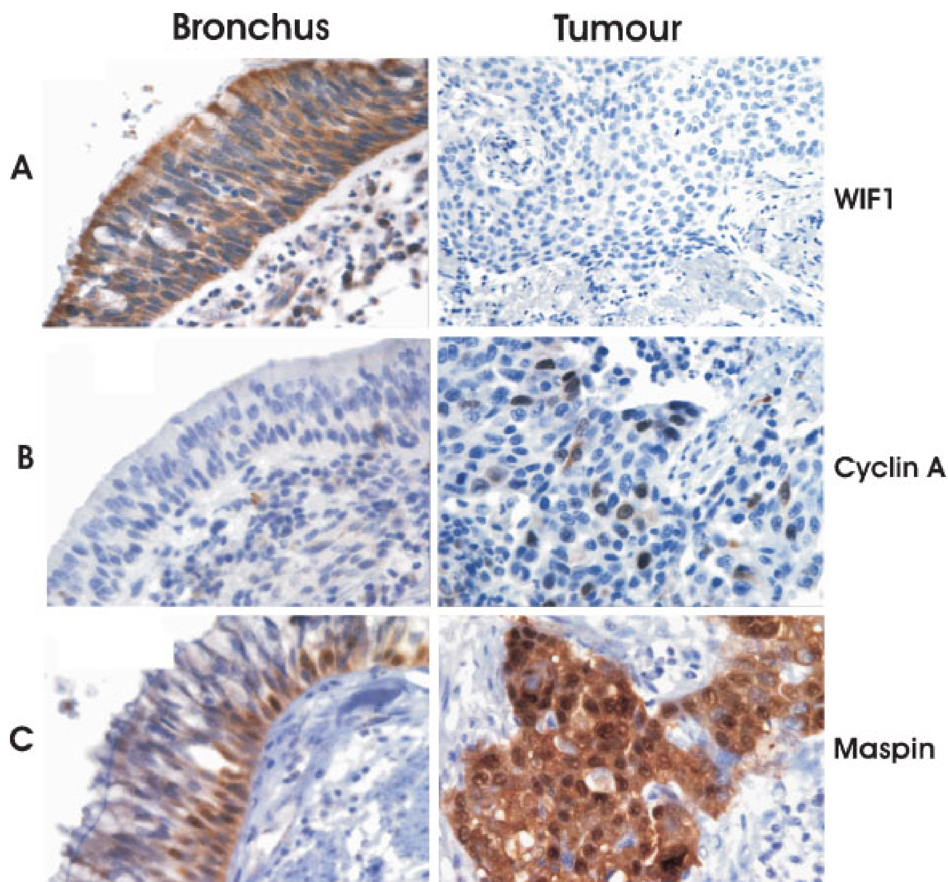

Figure 3. Immunoreactivity for WIF1, cyclin A, maspin, P-cadherin, FGFR3, MMP 7, and β-catenin in NSCLC and matched benign bronchial epithelium, corresponding to RNA expression levels: WIF1 (A) with strong reduction in squamous cell carcinoma. Cyclin A (B) with weak nuclear staining in normal bronchial epithelium and stronger expression in tumour cells. Maspin (C) with predominantly nuclear expression in basal cells of the bronchial epithelium and positive cytoplasmic and nuclear staining in cancer tissue. (D) P-cadherin with basal bronchial staining and scattered staining in carcinoma. (E) MMP 7 with negative bronchial staining and membranous cytoplasmic staining in carcinoma. (F) FGFR 3 in a predominantly basal location in bronchi and diffuse cytoplasmic staining in carcinoma. (G1) Strong cytoplasmic staining for β-catenin in bronchi and cancer tissue. Original magnification: $\times 200$, except WIF1 carcinoma and β-catenin bronchi, $\times 400$. (G2) Strong cytoplasmic staining bronchi of patients with smoking history and a weaker expression in the bronchi of non-smokers

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

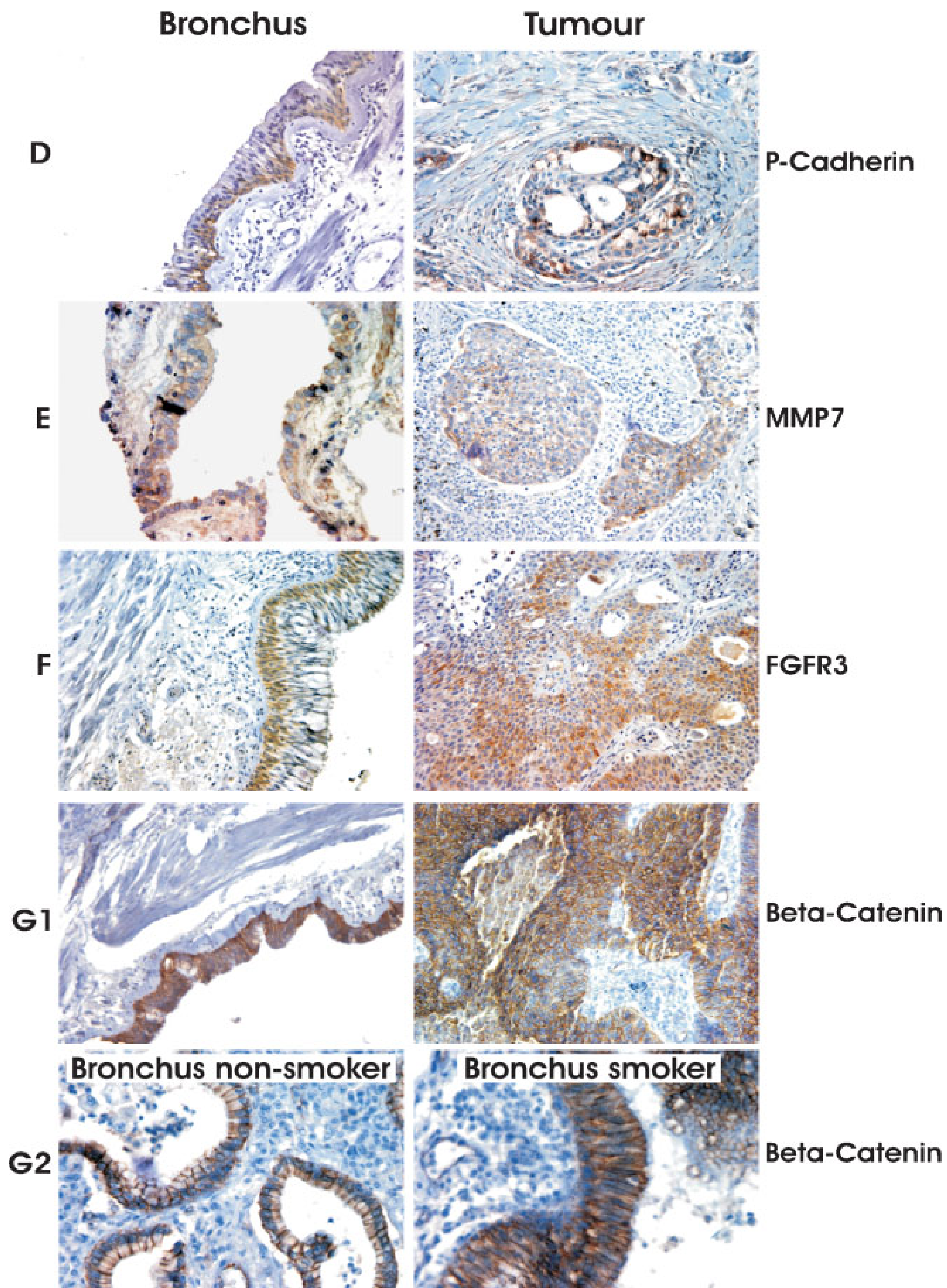

Figure 3. Continued

[5], who investigated the gene expression profiles of adenocarcinomas from smokers and non-smokers in a larger number of cases. Yatabe *et al* [23] described bronchioalveolar carcinomas affecting females and non-smokers that showed a mutation of EGFR. Three of our non-smokers with adenocarcinomas were female and two showed a bronchioalveolar growth pattern. All squamous cell carcinomas with a positive smoking history clustered in one subgroup, independent of the varying load of tobacco exposure [60 pack years (PY) on average in pool 1 and 31 PY in pool 4]. We found differentially expressed genes controlling cell proliferation in both squamous cell cancers and adenocarcinomas of smokers. These included *GMNN, CCNA2, Elf3*, and *CCNB1* [24–27] in squamous carcinomas, and *CDKN3* and *CCNG2* [26] in adenocarcinomas. Other genes play a role in signal transduction (*DDK3, WIF-1*) [19,28] and development (*SPRPB1, SPRPB2*) [29] in squamous cancers, and in adenocarcinomas (*MUC5, MUC13, MAGEA2, SFRP4, HOXA10* [30–32]). Genes involved in remodelling of the extracellular matrix included *SERPIN5* (maspin),

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

MMP11, MMP12, and *COL11A1* in squamous cell carcinomas, and *COL3A1, MMP7, MMP13, LAMB1*, and *COL 11 A* [33,34] in adenocarcinomas.

Relative quantitative real-time PCR was used to confirm the differential expression of eight selected genes which are involved in detoxification (*AKR1B10, ALDH3* [35,36]), tumour suppressor genes (*BLU, GAS2, MGP, MSM, WIF1* [37–39]), and are known as a neuroendocrine marker (*SCG2* [40]). Immunohistochemistry of six selected gene products that were differentially expressed at the RNA level found consistent results. These genes include an oncogene (FGFR3), a cell proliferation gene (cyclin A [24]) and matrix remodelling genes (β-catenin [41], wif-1 [19], maspin [33], MMP-7 [42], P-cadherin [43]). Although P-cadherin showed strong up-regulation at the RNA level, the immunohistochemical staining of tumour tissue was negative, which could be explained by post-transcriptional degradation. The other immunohistochemical staining results confirmed the findings of RNA expression.

The other major arm of our hierarchical cluster analysis contained only benign tissue samples and split into benign bronchial and peripheral alveolar tissue. While alveolar lung tissue showed no further segregation, the bronchial epithelium of smokers and non-smokers was arranged in two subgroups. Using the GOstat tool in the analysis of bronchial epithelium of current smokers and non-smokers, we found 30 significantly overrepresented genes in the categories cell adhesion and inflammation. Tobacco smoke-exposed bronchial samples revealed down-regulated differentially expressed genes that are involved in endothelial differentiation (eg CD31, thrombospondin 2), possibly representing chronic bronchitis. However, a direct effect of tobacco smoke on the expression of these genes cannot be excluded. Strongly up-regulated β-catenin and nuclear expression is found in many malignant tumours and is a sign of dysregulation of the wnt pathway [44]. These data suggest that smoking can influence the wnt pathway directly [41]. We performed cross-comparative analyses to find other putative molecular markers for cancer risk assessment by searching for consistent molecular alterations in bronchial epithelium and squamous cell carcinomas of smokers. Consistently up-regulated aldoketoreductase (*AKR1B10*) [35,45,46] and down-regulated flavin-containing monooxygenases (*FMO2, FMO3*) in tobacco smoke-exposed tissues mediate the metabolic activation of drugs and xenobiotic compounds to carcinogenic intermediates [47]. Eleven other genes were consistently down-regulated (ie *GLRX-1, ALDH1B, CYP26A1, HMGCS2, HLF, APIN, RTN1, LMO3, ITM2A, NTS, ANK2* [48–50]) and are involved in detoxification processes. Decreased expression of *ALDH1B* and *GLRX-1* is associated with multiple head and neck squamous cell cancers of smokers [51,52]. Up-regulation of *SPRP1B* (small proline-rich protein) is associated with early metaplastic squamous differentiation and lung carcinogenesis [29]. Results obtained in adenocarcinomas show superoxide dismutase 2 (*SOD2*), which is a member of the antioxidant defence system. *FGFR3* (fibroblast growth factor receptor 3) is an oncogene in many bladder cancers [53]. Because adenocarcinomas may derive from the alveolar part of the lung, we looked for consistent molecular alterations in the alveolar lung tissue of smokers. Glutathione peroxidase [54], which is involved in detoxification, is one example of a consistently up-regulated gene. The down-regulated probe set identified carboxypeptidase A3 [55], which regulates cancer-related pathways.

Since the present study included only a limited number of probes, further studies with validation experiments and downstream analyses are necessary and must validate these markers in sputum samples from high-risk smokers. Such gene signatures may be utilized to identify smokers at increased risk for lung cancer.

Acknowledgement

This work was supported by the Wilhelm-Sander-Foundation 2000.127.1.

Supplementary material

Supplementary material may be found at the web address http://www.interscience.wiley.com/jpages/0022-3417/suppmat.path.2039.html

References

1. Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics. *CA Cancer J Clin* 2003;**53:**5–26.
2. Kerr KM. Pulmonary preinvasive neoplasia. *J Clin Pathol* 2001;**54:**257–271.
3. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, *et al*. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2001;**98:**13784–13789.
4. Wigle DA, Jurisica I, Radulovich N, Pintilie M, Rossant J, Liu N, *et al*. Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. *Cancer Res* 2002;**62:**3005–3008.
5. Powell CA, Xu G, Filmus J, Busch S, Brody JS, Rothman PB. Oligonucleotide microarray analysis of lung adenocarcinoma in smokers and nonsmokers identifies GPC3 as a potential lung tumor suppressor. *Chest* 2002;**121:**6S–7S.
6. Petty RD, Nicolson MC, Kerr KM, Collie-Duguid E, Murray GI. Gene expression profiling in non-small cell lung cancer: from molecular mechanisms to clinical application. *Clin Cancer Res* 2004;**10:**3237–3248.
7. Wistuba II, Behrens C, Milchgrub S, Bryant D, Hung J, Minna JD, *et al*. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene* 1999;**18:**643–650.
8. Boyle JO, Lonardo F, Chang JH, Klimstra D, Rusch V, Dmitrovsky E. Multiple high-grade bronchial dysplasia and squamous cell carcinoma: concordant and discordant mutations. *Clin Cancer Res* 2001;**7:**259–266.
9. Gatenby RA, Vincent TL. An evolutionary model of carcinogenesis. *Cancer Res* 2003;**63:**6212–6220.

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

10. Gazdar AF, Minna JD. Molecular detection of early lung cancer. *J Natl Cancer Inst* 1999;**91:**299–301.
11. Gebel S, Gerstmayer B, Bosio A, Haussmann HJ, Van Miert E, Muller T. Gene expression profiling in respiratory tissues from rats exposed to mainstream cigarette smoke. *Carcinogenesis* 2004;**25:**169–178.
12. Murray JI, Whitfield ML, Trinklein ND, Myers RM, Brown PO, Botstein D. Diverse and specific gene expression responses to stresses in cultured human cells. *Mol Biol Cell* 2004;**15:** 2361–2374.
13. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, *et al*. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A* 2004;**101:**10143–10148.
14. Klimberg VS, Westbrook KC, Korourian S. Use of touch preps for diagnosis and evaluation of surgical margins in breast cancer. *Ann Surg Oncol* 1998;**5:**220–226.
15. Baugh LR, Hill AA, Brown EL, Hunter CP. Quantitative analysis of mRNA amplification by *in vitro* transcription. *Nucleic Acids Res* 2001;**29:**E29.
16. Beissbarth T, Speed TP. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* 2004;**20:**1464–1465.
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995;**57:**289–300.
18. Bettstetter M, Woenckhaus M, Wild PJ, Ruemmele P, Blaszyk H, Hartmann A, *et al*. Elevated nuclear maspin expression is associated with microsatellite instability and high tumor grade in colorectal cancer. *J Pathol* 2005;**205:**606–614.
19. Wissmann C, Wild PJ, Kaiser S, Roepcke S, Stoehr R, Woenckhaus M, *et al*. WIF1, a component of the Wnt pathway, is downregulated in prostate, breast, lung, and bladder cancer. *J Pathol* 2003;**201:**204–212.
20. Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, *et al*. Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am J Pathol* 1999;**154:**83–95.
21. Stoehr R, Wild P, Hartmann A. Lasermicrodissection — an important prerequisite for the molecular-genetic analysis of bladder cancer. *Pathol Res Pract* 2003;**199:**355–362.
22. Kendziorski CM, Zhang Y, Lan H, Attie AD. The efficiency of pooling mRNA in microarray experiments. *Biostatistics* 2003;**4:**465–477.
23. Yatabe Y, Kosaka T, Takahashi T, Mitsudomi T. EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. *Am J Surg Pathol* 2005;**29:**633–639.
24. Liu D, Matzuk MM, Sung WK, Guo Q, Wang P, Wolgemuth DJ. Cyclin A1 is required for meiosis in the male mouse. *Nature Genet* 1998;**20:**377–380.
25. Kulartz M, Knippers R. The replicative regulator protein geminin on chromatin in the HeLa cell cycle. *J Biol Chem* 2004;**279:**41686–41694.
26. Kopp JL, Wilder PJ, Desler M, Kim JH, Hou J, Nowling T, *et al*. Unique and selective effects of five Ets family members, Elf3, Ets1, Ets2, PEA3, and PU.1, on the promoter of the type II transforming growth factor-beta receptor gene. *J Biol Chem* 2004;**279:**19407–19420.
27. Chinami M, Yano Y, Yang X, Salahuddin S, Moriyama K, Shiroishi M, *et al*. Binding of HTm4 to cyclin-dependent kinase (Cdk)-associated phosphatase (KAP).Cdk2.cyclin A complex enhances the phosphatase activity of KAP, dissociates cyclin A, and facilitates KAP dephosphorylation of Cdk2. *J Biol Chem* 2005;**280:**17235–17242.
28. Hsieh SY, Hsieh PS, Chiu CT, Chen WY. Dickkopf-3/REIC functions as a suppressor gene of tumor growth. *Oncogene* 2004;**23:**9183–9189.
29. Lau D, Xue L, Hu R, Liaw T, Wu R, Reddy S. Expression and regulation of a molecular marker, SPR1, in multistep bronchial carcinogenesis. *Am J Respir Cell Mol Biol* 2000;**22:**92–96.
30. Lopez-Ferrer A, Curull V, Barranco C, Garrido M, Lloreta J, Real FX, *et al*. Mucins as differentiation markers in bronchial epithelium. Squamous cell carcinoma and adenocarcinoma display similar expression patterns. *Am J Respir Cell Mol Biol* 2001;**24:**22–29.
31. Zou H, Molina JR, Harrington JJ, Osborn NK, Klatt KK, Romero Y, *et al*. Aberrant methylation of secreted frizzled-related protein genes in esophageal adenocarcinoma and Barrett's esophagus. *Int J Cancer* 2004;**116:**584–591.
32. Lane D, Rutherford TJ, Taylor HS. HOXA10 expression in endometrial adenocarcinoma. *Tumour Biol* 2005;**25:**264–269.
33. Heighway J, Knapp T, Boyce L, Brennand S, Field JK, Betticher DC, *et al*. Expression profiling of primary non-small cell lung cancer for target identification. *Oncogene* 2001;**21:** 7749–7763.
34. Muller D, Breathnach R, Engelmann A, Millon R, Bronner G, Flesch H, *et al*. Expression of collagenase-related metalloproteinase genes in human lung or head and neck tumours. *Int J Cancer* 1991;**48:**550–556.
35. Hyndman D, Flynn TG. The aldo-keto reductases and their role in cancer. *Adv Exp Med Biol* 1999;**463:**427–434.
36. Sladek NE, Kollander R, Sreerama L, Kiang DT. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens. *Cancer Chemother Pharmacol* 2002;**49:**309–321.
37. Agathanggelou A, Dallol A, Zochbauer-Muller S, Morrissey C, Honorio S, Hesson L, *et al*. Epigenetic inactivation of the candidate 3p21.3 suppressor gene BLU in human cancers. *Oncogene* 2003;**22:**1580–1588.
38. Solomon B, Binns D, Roselt P, Weibe LI, McArthur GA, Cullinane C, *et al*. Modulation of intratumoral hypoxia by the epidermal growth factor receptor inhibitor gefitinib detected using small animal PET imaging. *Mol Cancer Ther* 2005;**4:**1417–1422.
39. Grepmeier U, Dietmaier W, Merk J, Wild PJ, Obermann EC, Pfeifer M, *et al*. Deletions at chromosome 2q and 12p are early and frequent molecular alterations in bronchial epithelium and NSCLC of long-term smokers. *Int J Oncol* 2005;**27:**481–488.
40. Clegg N, Ferguson C, True LD, Arnold H, Moorman A, Quinn JE, *et al*. Molecular characterization of prostatic small-cell neuroendocrine carcinoma. *Prostate* 2003;**55:**55–64.
41. Kato Y, Hirano T, Yoshida K, Yashima K, Akimoto S, Tsuji K, *et al*. Frequent loss of E-cadherin and/or catenins in intrabronchial lesions during carcinogenesis of the bronchial epithelium. *Lung Cancer* 2005;**48:**323–330.
42. Thomas P, Khokha R, Shepherd FA, Feld R, Tsao MS. Differential expression of matrix metalloproteinases and their inhibitors in non-small cell lung cancer. *J Pathol* 2000;**190:**150–156.
43. Charalabopoulos K, Gogali A, Kostoula OK, Constantopoulos SH. Cadherin superfamily of adhesion molecules in primary lung cancer. *Exp Oncol* 2004;**26:**256–260.
44. Choi YS, Shim YM, Kim SH, Son DS, Lee HS, Kim GY, *et al*. Prognostic significance of E-cadherin and beta-catenin in resected stage I non-small cell lung cancer. *Eur J Cardiothorac Surg* 2003;**24:**441–449.
45. Penning TM. AKR1B10: a new diagnostic marker of non-small cell lung carcinoma in smokers. *Clin Cancer Res* 2005;**11:**1687–1690.
46. Fukumoto S, Yamauchi N, Moriguchi H, Hippo Y, Watanabe A, Shibahara J, *et al*. Overexpression of the aldo-keto reductase family protein AKR1B10 is highly correlated with smokers' non-small cell lung carcinomas. *Clin Cancer Res* 2005;**11:**1776–1785.
47. Ryu SD, Yi HG, Cha YN, Kang JH, Kang JS, Jeon YC, *et al*. Flavin-containing monooxygenase activity can be inhibited by nitric oxide-mediated S-nitrosylation. *Life Sci* 2004;**75:**2559–2572.
48. Aledo R, Zschocke J, Pie J, Mir C, Fiesel S, Mayatepek E, *et al*. Genetic basis of mitochondrial HMG-CoA synthase deficiency. *Hum Genet* 2001;**109:**19–23.
49. Daheron L, Brizard F, Millot F, Cividin M, Lacotte L, Guilhot F, *et al*. E2A/HLF fusion gene in an acute lymphoblastic leukemia patient with disseminated intravascular coagulation and a normal karyotype. *Hematol J* 2002;**3:**153–156.
50. Aoyama M, Ozaki T, Inuzuka H, Tomotsune D, Hirato J, Okamoto Y, *et al*. LMO3 interacts with neuronal transcription

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

factor, HEN2, and acts as an oncogene in neuroblastoma. *Cancer Res* 2005;**65:**4587–4597.

51. Morita M, Oyama T, Kagawa N, Nakata S, Ono K, Sugaya M, *et al*. Expression of aldehyde dehydrogenase 2 in the normal esophageal epithelium and alcohol consumption in patients with esophageal cancer. *Front Biosci* 2005;**10:**2319–2324.
52. Padilla CA, Bajalica S, Lagercrantz J, Holmgren A. The gene for human glutaredoxin (GLRX) is localized to human chromosome 5q14. *Genomics* 1996;**32:**455–457.
53. van Rhijn BW, van Tilborg AA, Lurkin I, Bonaventure J, de Vries A, Thiery JP, *et al*. Novel fibroblast growth factor receptor 3 (FGFR3) mutations in bladder cancer previously identified in non-lethal skeletal disorders. *Eur J Hum Genet* 2002;**10:**819–824.
54. Moscow JA, Schmidt L, Ingram DT, Gnarra J, Johnson B, Cowan KH. Loss of heterozygosity of the human cytosolic glutathione peroxidase I gene in lung cancer. *Carcinogenesis* 1994;**15:**2769–2773.
55. Huang H, Reed CP, Zhang JS, Shridhar V, Wang L, Smith DI. Carboxypeptidase A3 (CPA3): a novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells. *Cancer Res* 1999;**59:**2981–2988.

Copyright © 2006 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.