PERGAMON

European Journal of Cancer 39 (2003) 1936–1947

European Journal of Cancer

www.ejconline.com

Gene expression profiles in human non-small and small-cell lung cancers

S. Difilippantonio[a,1,2], Y. Chen[a,1], A. Pietas[a], K. Schlüns[a], M. Pacyna-Gengelbach[a], N. Deutschmann[a], H.M. Padilla-Nash[b], T. Ried[b], I. Petersen[a,*]

[a]*Institute of Pathology, University Hospital Charité, Schumannsrt. 20/21 Berlin 10098, Germany*
[b]*Genetics Department, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA*

Received 17 October 2002; received in revised form 6 February 2003; accepted 13 February 2003

Abstract

Suppression subtractive hybridisation (SSH) was performed comparing normal bronchial epithelial cells with a lung squamous cell carcinoma (SCC) and a metastatic small-cell lung carcinoma (SCLC). The sequence analysis of four cDNA libraries revealed 869 individual sequences. Of these, 342 were tested using northern blots of lung cancer cell lines representing the three major subtypes (SCC, adenocarcinoma, SCLC) which confirmed the differential expression of 236 cDNAs. The extended analysis of 31 randomly chosen fragments confirmed the validity of the approach to identify genes associated with lung cancer development. Additionally, five novel full-length cDNA were isolated encoding the microtubule-associated proteins 1A/1B light chain 3, the epithelial V-like antigen 1 (*EVA1*), the GTP-binding protein *SAR1*, a new member of the S100-type calcium binding protein family and a new homeobox-containing gene.

Keywords: Gene expression profile; Suppression subtractive hybridisation; Human non-small cell lung cancer; Small-cell lung cancer; Sequencing; Gene mapping and full-length cDNA

1. Introduction

Lung cancer has a 5-year survival rate of approximately 15% and is a leading cause of cancer-related deaths. Although the incidence of this disease in men has decreased in recent years, the number of cases among women has increased significantly and is currently surpassing that of breast cancer. Lung cancer can be divided into two histological groups. Small-cell lung carcinoma (SCLC) comprises approximately 20% of the total incidence, while non-small cell lung carcinoma (NSCLC), which includes adenocarcinoma, squamous cell carcinoma (SCC) and large-cell carcinoma, accounts for 80% [1].

The origins of SCLC and NSCLC are still controversial. Since SCLC express neuroendocrine markers, they are often classified in the same category as pulmonary carcinoids and large cell neuroendocrine carcinomas. However, intermediate tumour stages between the mostly benign carcinoids and the aggressive SCLC, have only very rarely been identified, arguing against a common cell of origin. In contrast, SCLC frequently exist as an admixture with NSCLC [2]. These clinical and pathological features are validated by cytogenetic findings that show a significant overlap of chromosomal alterations in NSCLC and SCLC [3,4]. Recently, a study analysing expression profiles suggested that SCLC share more similarities with bronchial epithelial cells rather than pulmonary carcinoids [5]. These data suggest that both SCLC and NSCLC are derived from an epithelial precursor. Moreover, advanced SCC typically have chromosomal abnormalities similar to SCLC cells [6].

We previously isolated genes that were differentially expressed in a lung adenocarcinoma-derived cell line and small airway epithelial cells using suppression subtractive hybridisation (SSH) [7]. In this study, we compared SCLC and SCC cell lines with normal bronchial epithelial cells. Since each of these cell lines contained morphological and genetic aberrations that are typical of lung cancer, we endeavoured to identify lung cancer-associated genes and distinguish SCLC and NSCLC based on differences in their expression profiles.

* Corresponding author. Tel.: +49-30-450-536050; fax: +49-30-450-536110.
E-mail address: iver.petersen@charite.de (I. Petersen).
[1] Both authors contributed equally to this study.
[2] Current address: Experimental Immunology Branch, National Cancer Institute/National Institutes of Health, Bethesda, MD, USA.

0959-8049/03/$ - see front matter © 2003 Published by Elsevier Science Ltd.
doi:10.1016/S0959-8049(03)00419-2

2. Materials and methods

2.1. Cell lines and primary cell culture

Human small airway epithelial cells (SAEC) and human bronchial epithelial cells (HBEC), obtained from Clonetics (San Diego, CA, USA), were cultured in the recommended media, including growth factors and other supplements, and harvested subconfluently prior to 10 population doublings. The NSCLC-derived cell lines established in our laboratory (D51, D54, D97 and D117) were grown in Leibovitz 15 media, 10% (v/v) FCS, 1% (v/v) L-glutamine and additional supplements. Immortalised HBEC lines (H2078, H9442, H9609) and lung cancer cell lines (DMS-79, H125, H157, H187, H2030, H209, H2170, H2228, H226, H23, H332, H446, H526, H82, N417, SHP-77) were purchased from the American Type Culture Collection (ATCC) (Rockville, MD, USA). Lung cancer cell lines A427, A549, BEN, COLO-668, COLO-677, COLO 699, CPC-N and DV90 were purchased from the German Collection of Microorganisms and Cell cultures (Braunschweig, Germany). All cells were grown as recommended.

2.2. Cloning of cDNA fragments by SSH

Total and Poly(A)$^+$ RNA isolation, cDNA synthesis and subtraction were performed using the polymerase chain reaction (PCR)-select™ cDNA subtraction kit (Clontech, Palo Alto, CA, USA) as previously described in Ref. [7]. Both tester and driver cDNA were synthesised from HBEC, H2170 and H526 cells. Subtraction in both directions, comparing HBEC-H2170 and HBEC-H526, resulted in the cloning of four libraries. Individual transformants were isolated from white colonies on X-gal/isopropyl-beta-D-thio-galatopyranoside (IPTG) agar plates. To assess the quality of the libraries with respect to redundancy and specificity, 25 cDNA clones were randomly picked from each library and their sequence and differential expression determined by northern blotting analysis.

2.3. Sequence analysis

Sequencing reactions were performed with IRD-labelled M13 forward or reverse primers using the thermosequenase fluorescent-labelling cycle-sequencing kit according to the manufacturer's protocol (Amersham, Aylesbury, UK). Sequences were determined on a LICOR sequencer (MWG Biotech, Ebersberg, Germany). Final sequence homologuey searches were done against the GenBank (nr) and EST (dbEST) databases using the BLASTN algorithm at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST).

2.4. Expression analysis

Ten micrograms of total RNA were run on 1.2% (w/v) agarose gels in the presence of 2.2 M formaldehyde and transferred onto nylon membranes (Amersham) by capillary transfer. Plasmid inserts were amplified with nested primers (Clontech) and ^{32}P-labelled by random priming (Megaprime labelling, Amersham). Following 30 min prehybridisation with Expresshyb solution (Clontech), the denaturated probe was hybridised at 58 °C overnight. Blots were washed with 2×SSC/0.1% (v/v) sodium dodecyl sulphate (SDS) at room temperature, 0.1×squamous-cell carcinoma (SSC)/0.1% (v/v) SDS at 60 °C and exposed to X-ray film at −80 °C. Films were scanned using BIO-RAD Laboratories Imaging Densitometer Model GS-670 with 64 μm and an 8-bit resolution. Autoradiographic intensity was analysed using the Molecular Analyst Software Version 1.4. A rectangular frame was applied to the image of the blot around each individual signal. The expression level for each gene was quantified after the background correction. Sample loading differences were corrected by normalisation with *β-actin*. The expression levels listed in Table 1 were calculated in reference to HBEC. The values of 0.75 and 1.5 were used as the cut-off points for downregulation and upregulation, respectively. For the comparison of cell line to cell line, we calculated pair-wise Pearson correlation coefficients using all 330 normalised gene expression measurements.

2.5. Identification of full-length cDNA clones

In order to isolate full-length cDNA clones corresponding to seven partial cDNA clones with no significant homology to any known gene, PCR products (HBEC-43, HBECII-52, HBEC-234 and HBEC-73, HBECII-41, HBECII-2 and HBEC-326) were sent to the German Resource Center of the Human Genome Project (Heidelberg, Germany) for screening of cDNA libraries 441, 404, 408 and 719. Out of 62 clones, 26 were confirmed as positive by Southern blotting hybridisation. Clones DKFZp441H1211Q2 (HBEC-43), DKFZp441K016Q2 (HBECII-52), DKFZp441A199Q2 (HBEC-234), DKFZp404H1043Q2 (HBEC-73) and DKFZp404O119Q2 (HBEC-326) contained the longest insert determined by a *NotI*/*SalI*-restriction digest, which is in concordance with the transcript size estimated by northern blotting hybridisation. The final full-length cDNA was determined by sequencing.

2.6. Mapping of cDNA fragments by FISH and PCR

Fluorescent *in situ* hybridisation (FISH) probe labelling was performed by PCR (50 ng plasmid-DNA, 0.3 mM of each nested primer 1 and 2, 1.5 mM $MgCl_2$, 0.1 mM each deoxyadenosine triphosphate (dATP),

Table 1
Differentially expressed genes

Sequence identity (GenBank)	Level of downregulation				Sequence identity (GenBank)	Level of upregulation			
	Accession	D51	H2170	H526		Accession	D51	H2170	H526
Signal transduction and cell-cycle associated molecules									
Transforming growth factor-beta-induced 68 kD (*TGFBI*), (*BIGH3*)	NM_000358	1.7	5.4	8.4	grb7	AB008789	1	2.6	1
Amphiregulin (schwannoma-derived growth factor) (*AREG*)	NM_001657	2	2.4	2.9	Ras homologue enriched in brain 2 (*RHEB2*)	NM_005614	1	3.2	1
Novel FGF receptor mRNA	M64347	3.4	37.6	2.7	*c-erb-B-2*	X03363	2.1	6.7	1
Growth factor-inducible 2A9 gene/*Calcyclin*	M14300			2.8	Beta-subunit signal transducing proteins Gs/Gi	X04526	1	2.5	1
Death receptor 6 (*DR6*)	NM_014452		3	2.2	*Nras*-related gene	NM_007158	1	2.1	1
Caveolin 1 (*CAV1*)	AF125348	2	4.7	7.2	v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homologue (*KRAS2*)	NM_004985	2	1.6	2.3
Caveolin 2 (*CAV2*)	NM_001233	18	6	18.2	*c-kit* proto-oncogene	X06182	1	1	2
p66shc (*SHC*)	U73377		2.8	5	vav 1 oncogene (*VAV1*)	NM_005428	1	1	2.1
LIM domain kinase 2 (*LIMK2*)	D45906	3			Mus musculus ect2 oncogene (*Ect2*)	NM_007900	1.7	1.9	9.4
Cyclin-dependent kinase inhibitor *p21*	L25610	5.8	24	2.7	Rat GAP-associated protein (*p190*)	M94721	3	1	1
14-3-3 sigma protein	AF029082	5.5	6.7	8	Receptor tyrosine kinase-like orphan receptor 1 (*ROR1*)	NM_005012	1	1.7	2.3
Guanine nucleotide-binding regulatory protein (G) alpha-inhibitory-subunit	J03005	3	2	2	Transforming growth factor-beta (*tgf-beta*)	M60316	1	1	17
Cyclin 1 (*CYC1*)	AF135162	1.8	2.5	2.6	*Etk*/*Bmx* cytosolic tyrosine kinase	AF045459	1	1	5.3
					Phospholipase C, gamma (phosphatidylinositol-specific) (*PLCG2*)	NM_002661	2.1	2	14.9
					MacMarcks	X70326	3.2	4.2	6.8
					Protein phosphatase 2C (beta)	AJ005801	3.4	2.6	11.7
					Tumour necrosis factor, alpha-induced protein 2 (*TNFAIP2*)	NM_006291	7.2	1.8	6.3
					Retinoblastoma-binding protein 7 (*RBBP7*)	NM_002893	1.8	1.9	4.1
					Transmembrane protein Jagged 1 (*HJ1*)	AF028593	2	1	1
					CDC-like kinase1 (*CLK1*)	NM_004071	1.7	2.8	13
					Cyclin-dependent kinase 4 (*CDK4*)	NM_000075	3.1	3.5	4.7
Nuclear proteins (transcription factors, DNA processing enzymes)									
Programmed cell death 4 (*PDCD4*)	NM014456	12.9	12	33.9	*RAN*, member RAS oncogene family	NM_006325	1.7	1.6	2.9
Paired-like homeodomain transcription factor 1 (*PITX1*)	NM_002653	3.4	2.6	8.1	*RanBP1* (Ran-binding protein 1)	D38076	2.1	2.4	1.6
Homeo Box B2 (*HOXB2*)	NM_002145	2.7	2	2.3	DEK oncogene (DNA binding) (*DEK*)	NM_003472	1	1	3.3
Small nuclear RNA activating complex, polypeptide 1, 43 kD (*SNAPC1*)	NM_003082	1.6	1.8	2	*c-myb*	M15024	1	1	2.3
CCAAT/enhancer binding protein (*C*/*EBP*), delta (*CEBPD*)	NM_003651	7.5	<100	16	*N-myc* oncogene	M13228	1	1	12.3

(*continued on next page*)

Table 1 (*continued*)

Sequence identity (GenBank)	Level of downregulation				Sequence identity (GenBank)	Level of upregulation			
	Accession	D51	H2170	H526		Accession	D51	H2170	H526
Centromere protein-A (*CENP-A*)	NM_001809	4.1			Ro ribonucleoprotein-binding protein 1	AF114818	1.7	1	2.5
Karyopherin (importin) beta 2 (*KPNB2*)	NM_002270	17.8	4	13.9	DNA replication licensing factor (*huMCM2*)	D83987	2.8	1.6	14.9
CUSP	AF091627	6	4.8	7.6	Histone H1 transcriptional factor large subunit 2A	S74703	1.6	1.8	4.4
					Histone H2B	AJ223352	1	2.2	1
					Activating transcription factor 5 (*ATF5*)	NM_012068	3	2.1	6.3
					Skn-1a/*Epoc-1*/*Oct-11* POU transcription factor	AF162278	1	1	8.9
					small nuclear ribonucleoprotein polypeptide G (*SNRPG*)	NM_003096	1	1.9	1
					Replication factor C (activator 1) 5 (36.5 kD) (*RFC5*)	NM_007370	1.9	1.9	2.9
					Topoisomerase I	J03250	2	2.2	1
					Topoisomerase II alpha	AF071747	2.3	1	5.1
					Nuclear factor related to kappa B binding protein (*NFRKB*)	NM_006165	7.7	6.9	31
					Karyopherin alpha 2 (RAG cohort 1, importin alpha 1) (*KPNA2*)	NM_002266	5.3	5.5	8.6
					RNA-binding protein (autoantigenic) (*RALY*)	NM_007367	1	1.9	3
					High-mobility group (non-histone chromosomal) protein 1 (*HMG1*)	NM_002128	1	2	5.9
					Nucleolar phosphoprotein p130 (*P130*)	NM004741	1	1.6	1.8
RNA and protein processing, protein transport and protein degrading									
Heat shock protein 90	D87666	2.6	2.1		Putative spliceosome associated-protein	AF081788	1	5.2	2.5
NRD convertase	U64898	1.5	1.6	1.7	Splicing factor 3a, subunit 1, 120kD (*SF3A1*)	NM_005877	4.8	4.7	11
Cathepsin B (*CTSB*)	NM_001908			3	Putative translation initiation factor A121/SUI1	AF100737	1	1	1.9
Calpain, large polypeptide L2 (*CAPN2*) (calcium-activated neutral protease (*CANP*)	NM_001748	1.6		3.4	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 3 (*PSMD3*)	NM_002809	1.6	1.8	2.1
Serine protease inhibitor, Kunitz type 1 (*SPINT1*)	NM_003710	4.3	1.6	3.1	Proteasome (prosome, macropain) subunit, alpha type, 4 (*PSMA4*)	NM_002789	1	2	2.3
Serine proteinase inhibitor (*VAKTI*)	NM_006846	1.7	2	2	Heat shock 70 kD protein 1 (*HSPA1A*)	NM_005345	1	1	9.3
Squamous cell carcinoma antigen = serine protease inhibitor (*SCCA1*)	S66896	5.2	5.5	88	Sam68-like phosphotyrosine protein alpha (*SALP*)	AF051321	1	1	5.2
Protease inhibitor 5 (maspin) (*PI5*)	NM_002639	3.1	1.7	4.3	Cathepsin C (*CTSC*)	NM_001814	2	3.5	2
Plasminogen activator inhibitor-1 (*PAI-1*)	M16006	84	16.8	<100					
Plasminogen activator inhibitor-2 (*PAI-2*)	M24657			10.1					
Cystatin A (stefin A) (*CSTA*)	NM_005213	50	50	<100					

(*continued on next page*)

Table 1 (*continued*)

Sequence identity (GenBank)	Level of downregulation				Sequence identity (GenBank)	Level of upregulation			
	Accession	D51	H2170	H526		Accession	D51	H2170	H526
Metabolic enzymes, transporters, ion channels									
Calcium-activated chloride channel protein 3 (*CaCC3*)	AF127980	8.2	7.1	4	Thioredoxin reductase GRIM12	AF077367	2.8	2.6	2.9
Chloride channel, calcium activated, family member 4 (*CLCA4*)	NM_012128	1.6	1.6	2	Oxysterol binding protein (*OSBP*)	NM_002556	1	1	2.3
Glutathione peroxidase	NM_002084	1.7	2.8	2.5	Progesterone binding protein (*HPR6.6*)	NM_006667	2.1	3.2	4.4
Vacuolar H(+)-ATPase subunit (*ATP6GL*)	NM_004888	2.8		2.3	ATP-driven ion pump (*ATP1AL1*), ATPase, Na+/K+ transporting	NM_001676	1	1	4
Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (*PTGS2*)	NM_000963	8.3		<100	Glutathione S-transferase p1 (*GSTP1*)	NM_000852	1.6	1.7	4.7
Sialyltransferase (*STHM*)	NM_006456	2.8	33.1	10	Dihydropyrimidinase-like 3 (*DPYSL3*)	NM_001387	2.9	2.5	9
					Bleomycin hydrolase (*BLMH*)	NM_000386	1.6	3.3	1.7
					Methylene tetrahydrofolate dehydrogenase (NAD+dependent), Methenyltetrahydrofolate cyclohydrolase (*MTHFD2*)	NM_006636	6.5	10.7	11.4
					Collapsin response mediator protein 1 (*CRMP1*)	NM_001313	2.2	2	22
					Farnesyltransferase, CAAX box, alpha (*FNTA*)	NM_002027	1	1.9	4.2
					Dolichyl-phosphate beta-glucosyltransferase (*ALG5*)	AF102850	3.2	2.8	15.7
Cytoskeletal components, glycoproteins, epidermal differentiation complex, tumour associated antigens, extracellular proteins									
Keratin 5	NM_000424	3	5	9.1	Melanoma antigen, family A, 6 (*MAGEA6*)	NM_005363	1	3.4	7.8
Keratin 6 isoform K6a (*KRT6A*)	L42583	18.6	8.5	21.6	MAGE-9 antigen (*MAGE9*)	U10694	8.2	52	55
Keratinocyte lectin 14 (HKL-14)	U06643	53.5	17	<100	Tumour antigen (L6)	M90657	3.1	2.9	3.5
Keratin 15 (KRT15)	NM_002275	2.5	2.2	1.7	Paraneoplastic antigen MA1 (*PNMA1*)	NM_006029	2.1	2.3	4.2
Keratin 16 (KRT16A)	AF061812	4.1	4.5	8.5	NY-CO-1, serologically-defined colon cancer antigen 1 (*SDCCAG1*)	NM_004713	2.1	1.7	4.4
Keratin 17 (KRT17)	NM 000422	10	5.9	10.9	Epithelial glycoprotein (*EGP*)	M32306	3.3	6	7.3
Integrin, beta 1 (ITGB1)	NM_002211	3.3	2.9	2.3	NICE-3 protein	AJ243664	2.5	2.1	3.8
Integrin, alpha 3 (antigen CD49c)	NM 005501		3.6	4.6	pM5 protein (*PM5*)	NM_014287	3.1	2.2	3.6
Integrin, alpha 5 (fibronectin receptor, alpha polypeptide) (*ITGA5*)	NM_002205		5.7	5	Villin 2 (*ezrin*)	NM003379	1.9	1	5.3
Integrin alpha 6	NM000210	4.9	4.1	n.d.	Stromal cell-derived factor-2 (*SDF2*)	NM_006923	1	2.1	2.4
Cell adhesion molecule (*CD44*)	M59040	15.3	10	6.8					
Filamin A, alpha (actin-binding protein-280) (*FLNA*)	NM_001456		2.4	4.1					
Thymosin beta-4	M17733			3.7					

(*continued on next page*)

Table 1 (*continued*)

Sequence identity (GenBank)	Level of downregulation				Sequence identity (GenBank)	Level of upregulation			
	Accession	D51	H2170	H526		Accession	D51	H2170	H526
Thymosin beta-10	S54005	9.7	7.7	36.9					
Connexin 26 (*GJP2*)	M86849	4.3	4.1	5.3					
Desmoplakin (*DPI, DPII*)	NM_004415	4.4	2.5	3					
Desmocollin type 4	D17427	23.1	7.1	6.6					
Desmoglein 3 (*DSG3*)	NM_001944	3	2.4	2.4					
BH-protocadherin (*PCDH7*)	NM002589	2.4	3.4	2					
Hexabrachion (*HXB*)	M55618	2.5	8.3	11.5					
Fibronectin, *FN1*	M10905	3.3	57.1	14.8					
Laminin, alpha 3	NM_000227	11.4	5.7	5.5					
Laminin, beta 3 (*LAMB3*)	NM000228	3.3	4	5					
Laminin-5 beta3 chain	D37766	8.9	3.5	<100					
Bullous pemphigoid antigen (*BPGA1*)	NM_001723	19.7	5.3	11.1					
Gelsolin (amyloidosis, Finnish type) (*GSN*)	X04412	2	2	3.6					
Beta-galactoside-binding lectin	X14829		5	11.8					
Parathyroid hormone-like hormone (*PTHLH*)	NM_002820	3.3	4.2	7.4					
Syndecan-1 gene	Z48199	3.4	2.2	6.4					
sphingolipid activator protein 1	J03015	1.9	2.9	4					
Small proline-rich protein 3 (*SPRR3*)	NM_005416	22.5	10.2	13					
Tissue factor	M27436	6	7.6	5.2					
Epithelial protein lost in neoplasm alpha (*EPLIN*)	AF198455	2.9	1.8						
Moesin (*MSN*)	NM_002444		3.5	4.2					
Epithelial V-like antigen 1 (*EVA1*)	AF275945	2.1	3.1	3.3					
Thrombospondin 1	NM003246	3	1.9	1.8					
Others									
FLJ 23001 fis	AC007969	24.9	2.3	40.2	Glia maturation factor, beta (*GMFB*)	NM_004124	1	3.3	8.3
S100-type calcium binding protein A14	AY007220	11.6	6.8	45	Stathmin	X53305	1	1	4.1
Absent in melanoma 1 (*AIM1*)	U83115	3.5	1.8	1.8	Dynamin-like protein DYNIV-11 (*DYNIV-11*)	AF151685	1.6	1	4.6
Annexin A2 (*ANXA2*)	NM_004039	4.9	3.8	19.3	Epithelial membrane protein (*EMP1, CL-20*)	NM_001423	1	4.2	1
Annexin A8 (*ANXA8*)	NM_001630	10	7.6	37	Mesenchymal stem cell protein *DSC92*	AF242770	3.5	4.6	9.8
Myoferlin (*MYOF*)	AF182316	3.7	4.2	4.5	Nebulin	NM_004543	7.9	6.6	>100
S100 calcium-binding protein A8 (calgranulin A) (*S100A8*)	NM_002964	11.1	9.3	10	Seizure-related gene product 6	D29763	1	1.6	8.2
Pantophysin	S72481	1.7	3.8	5.7	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1 (*DDX1*)	NM_004939	2	1	5.6
FXYD domain-containing ion transport regulator 3 (*FXYD3*)	NM_005971	90	8.3	100	CGI-107 protein	AF151865	10.6	7.2	15.4
Adrenal gland protein AD-005	AF110778	2.8	8.4	2	CGI-150 protein	AF151908	1	1	1.8
Metallothionein from cadmium-treated cells	V00594	4.1	4.7	4.7	Neuroendocrine-specific protein-like protein 1 (*NSPL1*)	AF119297	3.7	2.6	7.3
Metallothionein 1L (*MT1L*)	NM_002450	26.7	10.3	8.1	Thyroid receptor interactor (*TRIP7*)	L40357	7.97.9	4.7	62.5

(*continued on next page*)

Table 1 (*continued*)

Sequence identity (GenBank)	Level of downregulation				Sequence identity (GenBank)	Level of upregulation			
	Accession	D51	H2170	H526		Accession	D51	H2170	H526
Interleukin 1-alpha	X02531	13.2	7.5	20	L-Plastin	L05492	3.8	17.5	15.7
Oncostatin M receptor (*OSMR*)	NM_003999	2.1	2	2	Interleukin 8 (*IL8*)	M28130	1	6.4	1
FK506-binding protein (*FKBP63*)	AF089745	1.8	4.7	5.8	RAD21 (*S. pombe*) homologue (*RAD21*)	NM_006265	1	2	3.7
PTD016 protein	AF100745	2.8	1.7	1.6	Tubulin, beta polypeptide (*TUBB*)	NM_001069	2.6	1	4.1
ERF-1	X79067		3.2	3	Hypothetical protein, clone YR-29	NM_014886	3.3	3.1	10.3
Tis11d	U07802	3.2	3.5	3.6	Hypothetical protein FLJ20647	NM_017918	1	1	10.7
Hypothetical protein *FLJ10788*	NM_018221	2.4	2.5	2.5	KIAA0042	D26361	1.8	2.3	2.6
cDNA FLJ20500 fis, clone KAT09159	AK000507	1.7	2.8	3.9	KIAA0069	D31885	1.9	1	3
cDNA: FLJ21504 fis, clone COL05662	AK025157		1.9	2	KIAA0203	NM_014781	1	1	7.4
cDNA clone FLJ10697	AK001559	1.9	3.4	1.8	cDNA FLJ10292 fis, clone NT2RM1000257	AK001154	1.7	1.6	2.8
cDNA DKFZp434A0225	AL137349			3.3	cDNA: FLJ20886 fis, clone ADKA03257	AK024539	1	2.2	9.8
Clone PRO0899 mRNA	AF116607	2.3	1.7	2	cDNA: FLJ22049 fis, clone HEP09444	AK025702	1	1	2.9
Clone DT1P1A7 mRNA	U92985		5.9	9.5	cDNA: FLJ21971 fis, clone HEP05790	AK025624	2.3	3.6	3.8
Clone hRPK.318_A_15,	AC005837	1.6	2.1	2.6	cDNA DKFZp586K1123	AL080216	1.6	4.2	2.7
DNA clone RP1-144C9	AL096774	2.2	2	2.1	cDNA DKFZp434C035	AL137633	1.8	1.8	4
DNA clone RP5-875K15	AL157952	2.1	1.9		cDNA DKFZp564C2163	AL117596	1	2	7.3
DNA clone RP11-264J4	AL138688	4.9	4.3	4.6	Clone 23732 mRNA	U79258	1.7	1	4.1
DNA clone bG279B7	AL078644	2.9	3.1	2	Clone DJ0820A21	AC006008	1	1	>100
DNA clone 300O13	AL079334	7	7.1		Clone RPCI11-0018K20	AC011198	1	3.6	1
DNA clone 380A1	Z97653	2.1	2.3	2.3	Cosmid R28784	AC005954	5.9	1	8.3
DNA clone 496H19	AF116607	9.5	4.9	52.3	BAC RPCI11-466O4, RP11-201E13	AC005297	4.6	4	25
DNA clone 434O14	AL022398	2	2		DNA clone RP4-809F4	AL022400	1	1	4.4
GTP-binding protein *SAR1*	AY008268	8.8	3.2	3.2	DNA clone RP4-569D19	AL022334	1	1	2.1
EST zk66b11.s1	AA043347	3.3	2.9		DNA clone RP5-1103G7	AL034548	2.4	1	>100
EST HSU51712	U51712	5.9	8	5	DNA clone RP1-317G22	AL133390	3	2.3	3.3
EST yx93g05.r1	N39842	13.1	2.7	2.4	DNA clone 1000E10	AL096773	1	8.3	1
EST 600943453T1	BE250411	5.6	10	10	EST tm78e07.x1	AI570600	1	1	>100
EST372676	AW960605	2.4	5	4.7	EST wd69c11.x1	AI692427	2.5	4.7	5.8
Sequence 4 from Patent WO9955858	AX011608	3	3	3.3	EST 601431837F1	BE891302	1.6	1.7	2.5
Sequence 28 from Patent WO9947655	AX017484	5	4	33	EST zu44b01.r1	AA477574	1	1	2.9
dbEST_Id 6635324	BF188819	6.9	3.8	8.7	EST 601577787F1	BE743861	10.8	5.8	22.9
dbEST_Id 6635323	BF188818	22.3	9.1	19.7	dbEST_Id 6635322	BF188817	1	1	>100
dbEST_Id 6751339	BF294064	3	2.6	3.6	dbEST_Id 6635320	BF188815	1	1	>100
					dbEST_Id 6635321	BF188816	1	1	8.7
					Clone 24800	AF070622	2.8	1	7.7
					dbEST_Id 6751337	BF294062	1.9	2.1	2.7

n.d. not determined

deoxycytidine triphosphate (dCTP), deoxyguanidine triphosphate (dGTP), 0.075 mM deoxythymidine triphosphate (dTTP) (Promega), 0.032 mM Biotin-16-deoxyuridine triphosphate (dUTP) (Roche) and 250 U Taq polymerase (Perkin-Elmer)). After 2 min of denaturation, 40 cycles (94 °C 30 s, 68 °C 15 s, 72 °C 1 min) were followed by a 10-min extension at 72 °C. PCR products were purified by spin-columns (QIAGEN) and the concentration estimated by agarose gel electrophoresis. 800 ng of PCR product, 1 μl Herring sperm DNA (Promega 10 mg/ml) and 20 μl *Cot1* DNA (Gibco 1 mg/ml) were ethanol precipitated, denatured and applied to denatured normal metaphase spreads. After hybridisation overnight (O/N) at 37 °C, the probes were detected by avidin-fluorescein isothiocyanate (FITC) according to the previously published protocol in Ref. [8]. Chromosomes were counterstained with 4,6-diamidino-2-phenylindole hydrochloride (DAPI) and embedded in 90% (v/v) glycerol containing phenylene diamine (Sigma Aldrich).

Mapping by PCR was either performed with chromosome-specific DNA or radiation hybrid panels (Stanford G3 Human/Hamster RH Panel, Research Genetics, USA). PCR reactions were performed in a 50-μl volume (10 ng template DNA, 15 pmol of the gene-specific primers and 2.0 mM deoxynucleotide triphosphates (dNTPs)). The DNA was denatured, amplified by 35 cycles (95 °C 60 s, annealing temperature (AT) 15 s, 72 °C 30 s) and a final extension for 5 min at 72 °C. The following primers and annealing temperatures were used: (a) HBEC-71, 5′-TAAAGTCACTCTCCATCATCT (F), 5′-GTTAGCCATCCATATGATTG (R), AT 55 °C; (b) HBEC-239, 5′-ACCTCAATTGCTGGAGTG (F), 5′-CTACAGATATGCCTGTCC (R), AT 55 °C; (c) HBEC-319, 5′-CCTCCATTCTCTTATCCAG (F), 5′-CCTCCATTCTCTTATCCAG (R), AT 52 °C; (d) HBEC-326, 5′-AAGCAGTGGCTTCACTGGA (F), 5′-CCTCCATTCTCTTATCCAG (R), AT 48 °C; (e) HBEC-398, 5′-GCACAGAGGCAGGAAGAACAA-3′ (F), 5′-TGGCACTGGCTTGGTGAGTCA (R), AT 60 °C. The radiation hybrid (RH) data were submitted to the Stanford Human Genome Center for Genome Research sequence tag site (STS) mapping server, and statistical analysis was performed using the radiation hybrid mapper (RHMAPPER) software package (http://www.shgc.stanford.edu/RH/rhserverformnew. html).

3. Results

3.1. cDNA library construction and northern blotting analysis

To identify genes associated with lung cancer, SSH was performed using cDNA synthesised from normal human bronchial epithelial cells, the squamous carcinoma cell line H2170 and the cell line H526 derived from a metastatic SCLC. cDNA libraries were generated from four separate subtractions (see Materials and methods) and 400–500 individual clones were isolated for each. We sequenced 1900 clones and due to redundancy identified 869 unique sequences. Based on the identity of known sequences and our interest in further examining those clones with no homology in the database, the expression status of 342 cDNA fragments was analysed in HBEC, H2170, H526 and the adeno-carcinoma cell line, D51, by northern blotting hybridisation. Representative northern blots are shown in Fig. 1. In total, 236 cDNAs (69%) were confirmed to be differentially expressed by this approach. These genes are listed in Table 1 according to their putative function/localisation with the upregulated ones in the right columns and the downregulated in the left columns. The expression levels of the tumour cell lines in relation to the normal HBEC cells were calculated for each gene, indicating whether it was similarly or differentially expressed between the two NSCLC and the SCLC cell line.

For 31 cDNA clones, we extended the expression analysis to small airway epithelial cells (SAEC), three immortalised HBEC and 28 lung cancer cell lines (Table 2). In general, downregulation of these genes was found not to be restricted to the lung carcinomas from which the libraries were generated, but expression was generally similar in other lung carcinoma cell lines. All genes, except those that are known to be differentiation-specific (e.g. *SPRR3*), were expressed at similar levels in both the normal primary cells examined (HBEC and SAEC). The clone H526-71 (dbEST_Id 6635322) was selectively expressed only in the SCLC cell line from which the SSH library was constructed. Interleukin 1-alpha was strongly expressed only in one NSCLC line, as well as in the normal cells.

Overall comparison of the gene expression profiles was calculated by the pair-wise correlation coefficient from one sample to another. The results are summarised in Table 3.

3.2. Isolation of ESTs and full-length cDNAs

Out of the 869 partial cDNA clones, seven full-length cDNA clones with no known gene homologuey in the database were isolated. Three of these corresponded to the same gene.

Clone DKFZp441H1211Q2 is 2110-bp long, has an open reading frame of 125 amino acids, and was identified as human microtubule-associated proteins 1A/1B light chain 3 (accession number AF303888), a structural protein involved in the filamentous cross-bridging between microtubules and other skeletal elements.

DKFZp441K016Q2 (2602-bp) has an open reading frame of 215 amino acids. It has been identified as human epithelial V-like antigen 1 (*EVA1*) (accession

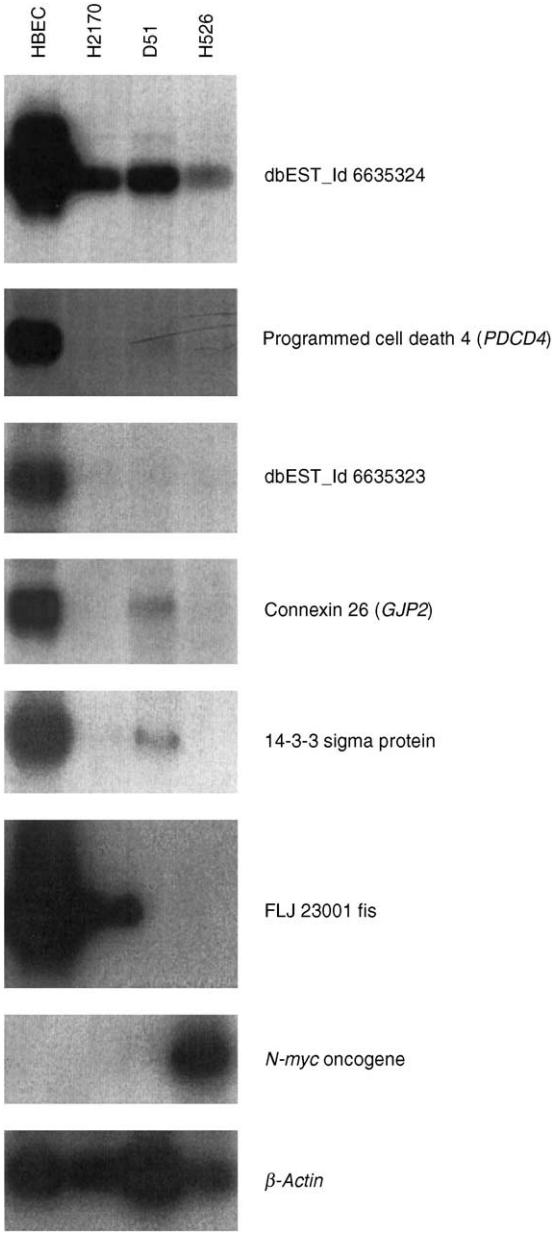

Fig. 1. Representative examples of northern blotting analysis using cloned cDNA fragments. 10 μg of total RNA from HBEC and the three lung cancer cell lines (H2170, D51 and H526) were electrophoresised, blotted and hybridised with radiolabelled cDNA probes and autoradiographed.

number AF304447) and may have a function in tumour development as a cell adhesion molecule.

DKFZp441A199Q2 (3003 bp) corresponds to the human GTP-binding protein SAR1 (accession number AY008268) and has an open reading frame of 198 amino acids.

DKFZp404H1043Q2 (1057 bp), with an open reading frame of 104 amino acids, is predicted to be a new member of S100-type calcium binding protein family (*S100A14*) (accession number AY007220). It was mapped to chromosome 1q21 by FISH [9].

DKFZp404O119Q2 (1014 bp), with a putative open reading frame of 73 amino acids, stands for the novel homeobox-containing gene *LAGY* (accession number AF454763). *LAGY* was mapped to chromosome 4 by PCR using chromosome-specific DNA together with radiation hybrid (RH) analysis. Meanwhile, we found that the gene is potentially involved in the development and progression of lung cancer [10].

Additionally, six out of all differentially expressed cDNA fragments showed no homology to any known genes or expressed sequence tags. They have been entered into the Database of Expressed Sequence Tags (dbEST) with the following designations: dbEST_Id 6635322 (accession number BF188817), dbEST_Id 6635320 (accession number BF188815), dbEST_Id 6635324 (accession number BF188819), dbEST_Id 6635323 (accession number BF188818), dbEST_Id 6635321 (accession number BF188816), dbEST_Id 6751337 (accession number BF294062).

3.3. Mapping of selected fragments

In addition to the *S100A14* gene and the *LAGY* gene, the following cDNA clones were mapped: cDNA FLJ13155fis (HBEC-398) to chromosome 3pter by FISH, FLJ 23001 fis (HBEC-239) to chromosome 2q23 by FISH and RH analysis, and finally dbEST_Id 6635323 (HBEC-71) to chromosome 4q by RH.

4. Discussion

SSH is a powerful approach for the identification of genes that are differentially expressed in one cell population compared with another. Although cDNA microarray is increasingly applied for the massively parallel analysis of gene expression, SSH is still widely used since it enables the recovery of abundant, as well as low-copy-number, mRNA transcripts. However, it not only permits an efficient identification of tumour-associated genes with known function, but also an unbiased isolation of novel sequences that are not yet available on the microchips. Using this approach, we compared normal human bronchial epithelial cells with a lung squamous carcinoma cell line and a metastatic SCLC.

We found that the proto-oncogene, *KIT*, was expressed solely in SCLC. This is in agreement with the postulated involvement of *KIT* in an autocrine loop in SCLC [11]. We also observed numerous other genes to be specifically expressed in SCLC including *MYB*, the *VAV 1* oncogene and the cytosolic tyrosine kinase *Etk*/*Bmx* (see Table 1).

Another group of genes identified was involved in cell-cycle regulation. Cyclin D1, p16, cdk4 and pRB are components of a regulatory pathway that is inactivated in most tumour cells [12]. Changes in the expression of

Table 2
Expression of selected lung cancer-associated genes

Name	Clones	Accession	Normal cells		Immortalized cells			Non-small cell lung cancer lines																		Smal- cell lung cancer lines									
			HBEC	SAEC	H2078	H9442	H9609	H2030	COLO699	DV90	D51	D54	D117	H23	A427	A549	H2228	H125	BEN	H157	H226	H2170	H322	D97	SHP77	CPC-N	H82	H446	Colo677	DMS79	H209	H187	N417	H526	COLO668
Tissue factor	HBEC-42	M27436	+++	+++					–		–	–	+	–				–	–		++	–		++	–	–	–	–	–	–				–	–
Interleukin 1-alpha	HBEC-131	X02531	+++	+++					–		–	–	–	–				–	–		+++	–		–	–	–	–	–	–	–				–	–
Clone 496H19	HBEC-26	AF116607	+++	+++	++	+	+		+		++	++	+	+	+		+	+	++	+	++	++	+	++	+	–	+	–	–	+	+		+		+	–
AIM1	HBEC-126	U83115	+++	+++					+		–	–	+	–	–			–	–	–	–	–	–	–	–	–	–	–	–	–				–	–
Clone DT1P1A7 mRNA	HBEC-11	U92985	+++	+++	+++	+				++	+++		+++	–	–	+	+					–	+++	+++	–	–	–		–	–			+	–	–
S100-type calcium binding protein A14	HBECII-20	AY007220	+++	+++	+++	++	+		–		–		–	–	–	–	–	–		–		+	++	–	–	–	–		–	–	–		+	–	–
DNA clone 300O13	HBEC-58	AL079334	+++	+++	+	+			–	++	–		–		+	–	+					+	+	–	–	+	++		–	–			++	++	+
Insulin-like growth factor binding protein 7 (IGFBP7)	HBEC-383	XM_071553	++	++				–			–	++	–					+	–		–	–		+	–		–		+	–				–	–
Integrin alpha 6	HBEC-90	NM000210	+++	+++				–			+	–	–					–	–		–	–		–	–		++		–	–				–	–
CUSP	HBEC-391	AF091627	+++	+++					–		–	–	–	–				–	–		–	–		–	–	–	–	–	–	–				–	–
Cell adhesion molecule (*CD44*)	HBEC-5	M59040	+++	+++				+			+	+++	+++					++	–		+++	–		+++	–		–		–	–				–	–
Laminin, alpha 3	HBEC-330	NM_000227	+++	+++				–			–	–	–					–	–		–	–		–	–		–		–	–				–	–
Integrin, beta 1 (*ITGB1*)	HBEC-139	NM_002211	+++	+++				+			+	+	+					+	+		+	+		+	+		+		+	+				+	+
Desmocollin type 4	HBEC-80	D17427	+++	+++					–		–	–	–	–	–			–	–	–	–	–	–	–	–	–	–	–	–	–				–	–
PAC 434O14	HBEC-81	AL022398	++	++	++	++	–		–		–	+	–	–	–		+	–		–	–	–	–	–	–	–	–	–	–	+	–		–	–	–
Epithelial membrane protein (*EMP1, CL-20*)	HBEC-290	NM_001423	–	–					–		–	–	+	–				+	–	++	++	++		++	–	–	–	–	–	–				–	–
Cystatin A (stefin A) (*CSTA*)	HBEC-60	NM_005213	+++	++				–			–	–	–					–	–		–	–		–	–		–		–	–				–	–
Nras-related gene	2170-333	NM_007158	+	+				+	+		+	+	+					+	+		+	+++		+	+		+		+	+				+	+
Moesin (*MSN*)	HBEC-155	NM_002444	++	++				+	+		++	++	++					++	+		++	–		++	+		+		++	–				–	++
Hypoxia-inducible factor 1	HBEC-118	XM_050771	++	++				+	–			–	++					–	–		++	–		++	+	++	++	–	–	–				+	+
DNA clone 88J8	HBEC-43	AL035402	++	++	++	–			–	–	–		–		–	–	–					–	–	–	–	–	–		–	–			–	–	–
Programmed cell death 4 (*PDCD4*)	HBEC-254	NM014456	+++	+++	+	–			–		–		–	–	–	–	–					–	–	–	–	+	+		–	+			+	–	–
cDNA FLJ13155fis	HBEC-398	AK023217	++	+	+	+	++				–			+	+		+	++				+	–	+	+	+			+	–	–		+	–	+
DNA clone bG279B7	HBEC-234	AL078644	+++	+++	+++	+++					+		+	++	+++		+++							+			+		+	+		–	+++	+	–
Epithelial V-like antigen 1 (*EVA1*)	HBEC-73	AF275945	++	++	–	–			–	–	–		–		–	–	–					–	–	–	–	–	–		–	–			+	–	–
SPRP3	HBEC-2	NM_005416	++	+	–	–	–		–	–	–		–	–	–	–	–	–		–		–	–	–	–	–	–		–	–	–	–	–	–	–
dbEST_Id 6635322	H526-71	BF188817	–	–			–				–		–	–	–		–	–			–	–	–	–	–	–	–		–	–			–	+	–
dbEST_Id 6635320	H526-85	BF188815	–	–			–				–		+	+	–		+	+			+	+	++	+	+	++	–		–	+			++	+++	–
dbEST_Id6751339	HBECII-302	BF294064	++	+			–				–		–	–	–		–	–			+	–		–	–	–	–		–	–			+		–
dbEST_Id6635324	HBEC-269	BF18819	+++	+++			–				–		–	–	–		–	–			+	–		–	+	–	–		–	+			–	+	–
novel homeobox-containing gene	HBEC-326	AF454763	+				–	–			–		–	–	–		–	–				–	–	–	–	–	–		–	–			–	–	–

Transcript levels: –, mRNA not detectable, +, ++, +++, low, intermediate or high mRNA expression.

Table 3
Pearson correlation coefficients by Pearson for pair-wise comparisons of samples

	HBEC	H2170	D51	H526
HBEC	1.000			
H2170	0.5222 ($P<0.01$)	1.000		
D51	0.478 ($P<0.01$)	0.775 ($P<0.01$)	1.000	
H526	0.113 ($P<0.05$)	0.526 ($P<0.01$)	0.571 ($P<0.01$)	1.000

these genes, like overexpression of CDK4, are predicted to lead to the loss of pRB function and therefore loss of cellular proliferative control. Furthermore, we identified the overexpression of the retinoblastoma-binding protein 7 in tumour cells. Although the relevance of this is unclear, overexpression of another protein, RBP1L1, from this family was considered as a diagnostic biological marker for a broad range of human cancers, including lung cancer [13]. In contrast, the p53-regulated cell-cycle inhibitors, p21 and 14-3-3 protein σ, were found to be downregulated in all of the cell lines analysed.

Growth arrest, increased expression of differentiation-associated proteins and morphological changes, are mediated by elevated extracellular calcium levels [14,15]. The loss of potential of lung tumour cells to undergo terminal differentiation may be a reflection of the transcriptional downregulation of calcium-binding proteins responsible for transmitting theses signals, like Calgranulin A (S100A8), CaN19 (S100A2), and probably the novel S100-type calcium binding protein, S100A14. Moreover, it comprises also the differentially expressed keratins and cornified envelope precursor proteins. The strong downregulation, or even absence, of cytokeratins K5, K6A, K15, K16, K17 and keratinocyte-lectin 14, as well as the expression of SPRR3, only in bronchial epithelial cells are consistent with studies on SCC of the head and neck and skin [16,17].

Another important event during the process of cancer invasion and metastasis is degradation of the extracellular matrix. The plasminogen activator/plasmin system of serine proteases has been implicated in these processes. Plasminogen activator was observed to be highly expressed in lung cell lines derived from metastatic patients and significantly elevated in prostate cancer patients with metastasis [18,19]. Our finding that the downregulation of plasminogen activator inhibitors 1 and 2 in lung tumour cell lines raises the speculation that they might inhibit the process of lung tumour invasion. Other serpins, such as maspin and Kunitz-type inhibitors, have been implicated in the suppression of prostate and breast cancer invasion [20,21]. Cystatin A (stefin A) has been inversely correlated with tumour relapse and the mortality rate in HNSCC [22].

Overall, several genes found to be differentially expressed in our study have also been described in HNSCC and breast carcinomas [23–25], thereby supporting the idea that malignancies of epithelial cells have common cancer-related expression profiles. Moreover, a significant proportion of the genes belong to the Ras-pathway or are p53-regulated genes [26,27]. This is not surprising since mutations of *Ras* and *p53* are known to play a major role in lung cancer tumourigenesis [28]. The utility of the applied method to assess functional changes occurring during the development and progression of lung cancer is supported by both the types of genes identified as differentially expressed and their biological function.

Interestingly, a number of the identified genes cluster within regions that are frequently affected by chromosomal imbalances. For example, the gene programmed cell death 4 (*PDCD4*) maps to chromosome 10q24, and could be a target of the deletions frequently observed in advanced lung carcinomas [29]. Moreover, the mapping of unknown or not yet characterised cDNA fragments reveals a useful application for the further identification of candidate genes within these regions.

Although concerns have been raised that cell lines may undergo unpredictable alterations *in vitro* which might result in diverse geno- and phenotypes, we think they comprise a useful model system for the identification of lung-cancer associated genes. Firstly, they facilitate the isolation of high quality mRNA, and they also allow for the recovery of carcinoma-cell-specific genes (without contamination by connective tissue, blood vessels and inflammatory cells that infiltrate the tumour mass) via a direct comparison between carcinoma cells and their epithelial precursors. Second of all, the lung carcinoma cells used in our study show chromosomal aberrations characteristic of their tumour types. Thirdly, among the 236 differentially expressed genes listed in Table 1, 123 genes were investigated in the 24 000-element cDNA microarray, and 39 were found in the subset of (918) clones that we identified to be most valuable for the clustering/classification of primary lung carcinomas using cDNA microarrays [30].

In this microarray study, the genes *keratin 17* and *annexin 8* were identified as class discriminators of lung SCC. These genes appeared to be downregulated in the H2170 lung SCC cell line compared with HBEC cells. However, the expression levels were relatively higher than in the adenocarcinoma cell line D51 and the SCLC cell line H526. Thus, these results are consistent with a role for these genes as markers for lung cancer differentiation. Our data also show that the observed expression differences are not unique to individual cell lines, since all 28 lung cancer cell lines examined displayed similar patterns. Although our correlation coefficient between normal bronchial epithelia (HBEC) and SCLC (H526) was lower than in a previous report, in which expression data was recovered from a 18 000 cDNA array [5], our method is designed to specifically generate expression differences, thus a greater deviation from 1.0 would be anticipated. The pair-wise correlation coefficient

analysis between individual NSCLC and SCLC cell lines was also significant. This fact, together with the data showing that SCLC are distinct from pulmonary carcinoids, suggests that both NSCLC and SCLC derive from a common epithelial precursor and underlines the need for a new classification of lung cancer that includes genetic composition and gene expression profiles.

In summary, we have provided detailed sequence information on transcriptional changes associated with lung carcinogenesis, with many of them being described for the first time. The specific set of genes we recovered by cDNA subtraction provides a basis for identifying transcripts with diagnostic and prognostic value, with respect to both primary tumours and metastases. In addition, studying the function of these genes and their biological pathways may lead to the development of new therapeutic options.

Acknowledgements

This work was supported by the Deutsche Krebshilfe (grant 10-1300-Pe2), the Deutsche Forschungsgesellschaft (DFG Pe602/1-2) and the Monika-Kutzner-Stiftung. The technical assistance of Cordula Heckert is gratefully acknowledged. We thank Dr. Michael J. Difilippantonio for helpful suggestions and comments during the preparation of the manuscript.

References

1. Pass HI, Mitchell JB, Johnson DH, Turrisi AT. *Lung Cancer: Principles and Practice*. Philadelphia, PA, Lippincott-Raven, 1996.
2. Roggli VL, Vollmer RT, Greenberg SD, McGavran MH, Spjut HJ, Yesner R. Lung cancer heterogeneity: a blinded and randomized study of 100 consecutive cases. *Hum Pathol* 1985, **16**, 569–579.
3. Petersen I, Langreck H, Wolf G, *et al*. Small-cell lung cancer is characterized by a high incidence of deletions on chromosomes 3p, 4q, 5q, 10q, 13q and 17p. *Br J Cancer* 1997, **75**, 79–86.
4. Petersen I, Bujard M, Petersen S, *et al*. Patterns of chromosomal imbalances in adenocarcinoma and squamous cell carcinoma of the lung. *Cancer Res* 1997, **57**, 2331–2335.
5. Anbazhagan R, Tihan T, Bornman DM, *et al*. Classification of small cell lung cancer and pulmonary carcinoid by gene expression profiles. *Cancer Res* 1999, **59**, 5119–5122.
6. Petersen S, Aninat-Meyer M, Schluns K, Gellert K, Dietel M, Petersen I. Chromosomal alterations in the clonal evolution to the metastatic stage of squamous cell carcinomas of the lung. *Br J Cancer* 2000, **82**, 65–73.
7. Petersen S, Heckert C, Rudolf J, *et al*. Gene expression profiling of advanced lung cancer. *Int J Cancer* 2000, **86**, 512–517.
8. Schriml LM, Padilla-Nash HM, Coleman A, *et al*. Tyramide signal amplification (TSA)-FISH applied to mapping PCR-labeled probes less than 1 kb in size. *Biotechniques* 1999, **27**, 608–613.
9. Pietas A, Schluns K, Marenholz I, Schaefer BW, Heizmann CW, Petersen I. Molecular cloning and characterization of the human S100A14 gene encoding a novel member of the S100 family. *Genomics* 2002, **79**, 513–522.
10. Chen Y, Petersen S, Pacyna-Gengelbach M, Pietas A, Petersen I. *Oncology* 2003, **64**, 450–458.
11. Krystal GW, Hines SJ, Organ CP. Autocrine growth of small cell lung cancer mediated by coexpression of c-kit and stem cell factor. *Cancer Res* 1996, **56**, 370–376.
12. Sherr CJ. Cancer cell cycles. *Science* 1996, **274**, 1672–1677.
13. Cao J, Gao T, Stanbridge EJ, Irie R. RBP1L1, a retinoblastoma-binding protein-related gene encoding an antigenic epitope abundantly expressed in human carcinomas and normal testis. *J Natl Cancer Inst* 2001, **93**, 1159–1165.
14. Hennings H, Michael D, Cheng C, Steinert P, Holbrook K, Yuspa SH. Calcium regulation of growth and differentiation of mouse epidermal cells in culture. *Cell* 1980, **19**, 245–254.
15. Yuspa SH, Kilkenny AE, Steinert PM, Roop DR. Expression of murine epidermal differentiation markers is tightly regulated by restricted extracellular calcium concentrations in vitro. *J Cell Biol* 1989, **109**, 1207–1217.
16. Leethanakul C, Patel V, Gillespie J, *et al*. Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* 2000, **19**, 3220–3224.
17. Lohman FP, Medema JK, Gibbs S, Ponec M, van de Putte P, Backendorf C. Expression of the SPRR cornification genes is differentially affected by carcinogenic transformation. *Exp Cell Res* 1997, **231**, 141–148.
18. Chen JJW, Peck K, Hong TM, *et al*. Global analysis of gene expression in invasion by a lung cancer model. *Cancer Res* 2001, **61**, 5223–5230.
19. Miyake H, Hara I, Yamanaka K, Arakawa S, Kamidono S. Elevation of urokinase-type plasminogen activator and its receptor densities as new predictors of disease progression and prognosis in men with prostate cancer. *Int J Oncol* 1999, **14**, 535–541.
20. Sheng S, Carey J, Seftor EA, Dias L, Hendrix MJ. Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells. *Proc Natl Acad Sci USA* 1996, **93**, 11669–11674.
21. Kobayashi H, Shinohara H, Takeuchi K, *et al*. Inhibition of the soluble and the tumour cell receptor-bound plasmin by urinary trypsin inhibitor and subsequent effects on tumour cell invasion and metastasis. *Cancer Res* 1994, **54**, 844–849.
22. Strojan P, Budihna M, Smid L, *et al*. Prognostic significance of cysteine proteinases cathepsins B and L and their endogenous inhibitors stefins A and B in patients with squamous cell carcinoma of the head and neck. *Clin Cancer Res* 2000, **6**, 1052–1062.
23. Fung LF, Lo AK, Yuen PW, Liu Y, Wang XH, Tsao SW. Differential gene expression in nasopharyngeal carcinoma cells. *Life Sci* 2000, **67**, 923–936.
24. Zhang M, Martin KJ, Sheng S, Sager R. Expression genetics: a different approach to cancer diagnosis and prognosis. *Trends Biotechnol* 1998, **16**, 66–71.
25. Nacht M, Ferguson AT, Zhang W, *et al*. Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer. *Cancer Res* 1999, **59**, 5464–5470.
26. Zuber J, Tchernitsa OI, Hinzmann B, *et al*. A genome-wide survey of RAS transformation targets. *Nat Genet* 2000, **24**, 144–152.
27. Zhao R, Gish K, Murphy M, *et al*. Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev* 2000, **14**, 981–993.
28. Sekido Y, Fong KM, Minna JD. Progress in understanding the molecular pathogenesis of human lung cancer. *Biochim Biophys Acta* 1998, **1378**, F21–F59.
29. Petersen S, Rudolf J, Bockmuhl U, *et al*. Distinct regions of allelic imbalance on chromosome 10q22-q26 in squamous cell carcinomas of the lung. *Oncogene* 1998, **7**, 449–454.
30. Garber ME, Troyanskaya OG, Schluns K, *et al*. Diversity of gene expression in adenocarcinoma of the lung. *Natl Acad Sci USA* 2001, **98**, 13784–13789.