SPANDIDOS PUBLICATIONS

# Identification and classification of differentially expressed genes in non-small cell lung cancer by expression profiling on a global human 59.620-element oligonucleotide array

HANS-STEFAN HOFMANN[1], BABETT BARTLING[1], ANDREAS SIMM[1], RICH MURRAY[2], NATASHA AZIZ[2], GESINE HANSEN[3], ROLF-EDGAR SILBER[1] and STEFAN BURDACH[4]

[1]Department of Cardio-Thoracic Surgery, Martin-Luther-University Halle-Wittenberg, Ernst-Grube-Str. 40, D-06097 Halle, Germany; [2]Eos Biotechnology Inc., Bioinformatics Department, Fremont, CA, USA; [3]Department of Pediatric Pulmonology and Neonatology, Hannover Medical School, Carl-Neuberg-Str. 1, D-30625 Hannover; [4]Department of Paediatrics and Comprehensive Cancer Center, Munich University of Technology, Kölner Platz 1, D-80804 Munich, Germany

Received February 9, 2006; Accepted April 4, 2006

**Abstract.** Improvements in detection, treatment and prognosis for patients with non-small cell lung cancer (NSCLC) depend on the molecular understanding of tumor development and progression. Using Affymetrix GeneChips comprising 59,620 elements, we determined the gene expression profiles of 89 NSCLC and 15 normal lung samples. We found 187 (0.3%) genes, which are at least 2-fold overexpressed and 157 (0.3%) genes 2-fold less expressed in NSCLC compared with normal lung. Cell cycle regulation, cell adhesion and nucleotide metabolism were the major biological processes connected to a large proportion of genes up-regulated in NSCLC. Down-regulated genes were frequently involved in metabolic/catabolic processes and signal transduction. The expression of specific genes revealed reliable differentiation of tumor from normal lung tissues, as well as the classification of both NSCLC subtypes squamous cell carcinoma and adenocarcinoma. In this context, collagens (COL7, 17) and cytokeratins (CK6, 15, 17) are preferentially induced in squamous cell carcinoma, whereas several transcription factors (TTF1, DAT1, TF-2) are exclusively elevated in adenocarcinomas. Some gene products involved in the metastatic process [matrixmetalloproteinase 12 (MMP-12) and urokinase plasminogen activator α (uPA)] were found as prognostic markers for the recurrence free interval and survival. Particularly, the simultaneous use of the MMP-12 and uPA expression predicted relapse-free and global survival of the patients. Screening of NSCLC with a genome-wide array revealed diagnostic, prognostic and potential therapeutic targets that might be suitable for an individual risk profile by tumor specific arrays.

*Correspondence to*: Dr Hans-Stefan Hofmann, Department of Cardiothoracic Surgery, Martin-Luther-University Halle-Wittenberg, Ernst-Grube-Str. 40, D-06097 Halle, Germany
E-mail: stefan.hofmann@medizin.uni-halle.de

*Key words:* non-small cell lung cancer, DNA-array

## Introduction

Lung cancer is the leading cause of cancer death worldwide in men and in women. As 86% of the people who are diagnosed with lung cancer die within 5 years, the 169,000 new cases in the United States will lead to more deaths than breast, colon, prostate and cervical cancers combined (1). The pathogenesis of lung cancer remains highly elusive due to its aggressive biologic nature and considerably heterogeneity, as compared to other cancers. Many genes and signaling pathways are known to be involved in the carcinogenesis as well as progression of non-small cell lung cancer (NSCLC). Genes potentially involved in NSCLC include the tumor suppressor gene p53 (2,3), the vascular endothelial growth factor receptor (EGFR) (4), the proliferation factor Ki-67 (5) and the anti-apoptotic Bcl-2 (6). However, none of these molecular markers have achieved a general use in diagnosis or prognosis of NSCLC. Up to now only tumor stage, determined by tumor size and extension, lymph node involvement and distant metastases have gained a widespread acceptance in the decision of therapy and prognosis of patients with NSCLC. Tumor development is associated with molecular changes and it is likely that many of the involved genes are currently unknown.

Global DNA-arrays are a high-throughput method to create an overview of the gene expression status of various malignancies and identify typical targets of these diseases (7). In this context, global human oligonucleotide arrays are also suitable to recognize differentially expressed genes in NSCLC. Therefore, the present study was performed to assess differentially expressed genes in NSCLC by DNA-microarray technology and their use for tissue discrimination and survival prognosis. The aim was the evaluation of target genes,

which may provide candidates for a lung cancer specific DNA-array.

## Materials and methods

*Patients and samples*. Tumor and normal lung tissue samples were obtained from 89 consecutive patients with NSCLC, who underwent pulmonary resection surgery between 1999 and 2001 (Table I). Only patients with clear histological classification as NSCLC (adenocarcinoma, squamous cell carcinoma) without neoadjuvant chemo- or radiotherapy were admitted to the study. In the case of curative resection of the tumor, an extended lymph node dissection was performed. Immediately following resection the tumor tissue and matched normal lung tissue were snap-frozen and stored in liquid nitrogen until use. Tumor histology and stages were estimated according to WHO (8) and the TNM staging according to the UICC classification, respectively (9). Clinical and pathological data were obtained for all patients and rendered anonymous. The median follow-up duration was 30.9 months (range 14-56 months).

*RNA preparation*. Total-RNA was prepared from cryopreserved lung cancer samples or normal lung tissue with acid phenol/chloroform extraction (TRIzol; Invitrogen, Karlsruhe, Germany) followed by purification with RNeasy Mini Kit (Qiagen; Hilden, Germany) according to manufacturer's instructions and quantified spectrophotometrically. RNA quality was assessed by visualization of 18S and 28S RNA bands after agarose gel electrophoresis and staining with ethidium bromide.

*Microarray expression analysis*. A total of 10 $\mu$g RNA from each sample was used to prepare biotinylated target cRNA as previously described (10,11). A detailed protocol is available at www.affymetrix.com. Samples were hybridized to a custom expression monitoring RNA microarray designed by Eos Biotechnology, Inc. using Affymetrix GeneChip technology (12) that contained essentially all expressed human genes in the public domain at the time of design (EOS-K). Sequences included on the array were derived from human genomic, expressed mRNA and EST databases in GenBank (13). Consensus sequences representing human expressed sequences were generated using the Clustering and Alignment Tool software (DoubleTwist, Oakland, CA), and prediction of the expressed genome from the human genome sequence was done using Ab intron-exon prediction (14). The 59,000 probesets on this microarray represent approximately 45,000 mRNA and EST clusters and 6,200 predicted exons. Data were used after $\gamma$-distribution normalization.

*Gene chip analysis*. After fragmentation of biotinylated cRNAs and hybridization with microarrays (EOS-K), signals were detected with streptavidin-phycoerythrin (SAPE). Signal enhancement was performed using biotinylated goat-anti-streptavidin antibodies. Arrays were washed and stained with the GeneChip Fluidics Station 400 and scanned with a GeneArray Scanner. Primary image analysis was performed using Microarray Suite 5.0. Images were scaled to an average hybridization intensity of 200. In total 49 squamous cell carcinomas, 40 adenocarcinomas and 15 corresponding normal lung samples were analyzed. All expression values below 60 were set to 60. To identify specific genes that were differentially expressed in tumors as compared to normal lung tissue we used a criterion that marks differential gene expression at an approximate significance level (determined by Bonferoni method) of $8.0 \times 10^{-7}$ using Student's t-test and a fold-change cut-off of 2.0 for up regulated, respectively 0.5 for down-regulated genes. Calculation of fold-changes was performed by dividing the mean expression level of a gene in the tumor samples by the mean expression level of the same gene in the lung samples.

Table I. Clinical and pathological characteristics of patients and their tumors.

| | | |
|---|---|---|
| No. of patient | 89 | |
| Mean age (yr) | 65.5 | |
| Male/female ratio | 71/18 | |
| Smoking/non-smoking ratio | 60/29 | |
| Surgical procedure | | |
| Probe excision/diagnostic thoracotomy | 6 | (6.8%) |
| Segment-resection | 5 | (5.5%) |
| Lobectomy | 64 | (72%) |
| Bilobectomy | 2 | (2.2%) |
| Pneumonectomy | 12 | (13.5%) |
| Histology | | |
| Squamous cell carcinoma | 49 | (55%) |
| Adenocarcinoma | 40 | (45%) |
| TNM staging | | |
| I | 31 | (34.8%) |
| II | 23 | (25.8%) |
| III | 29 | (32.6%) |
| IV | 6 | (6.8%) |
| Grading | | |
| Well/moderately well differentiated | 25 | (28%) |
| Poorly or undifferentiated | 64 | (72%) |
| Residual tumor situation | | |
| R0 | 75 | (84.2%) |
| R1 | 6 | (6.8%) |
| R2 | 8 | (9%) |

*Gene ontology classification*. Gene products of differently expressed genes were classified by the Gene Ontology (GO) classification to determine which biological functions or pathways were changed in the tumor. The GO classification consists of three structured, controlled vocabularies that describe gene products in terms of their associated biological processes, cellular component and molecular functions. These data are available at www.geneontology.org/.

*Statistical analysis*. For analysis of follow-up data, life table curves were calculated with Kaplan-Meier methods, and survival distributions were compared by log-rank statistics.

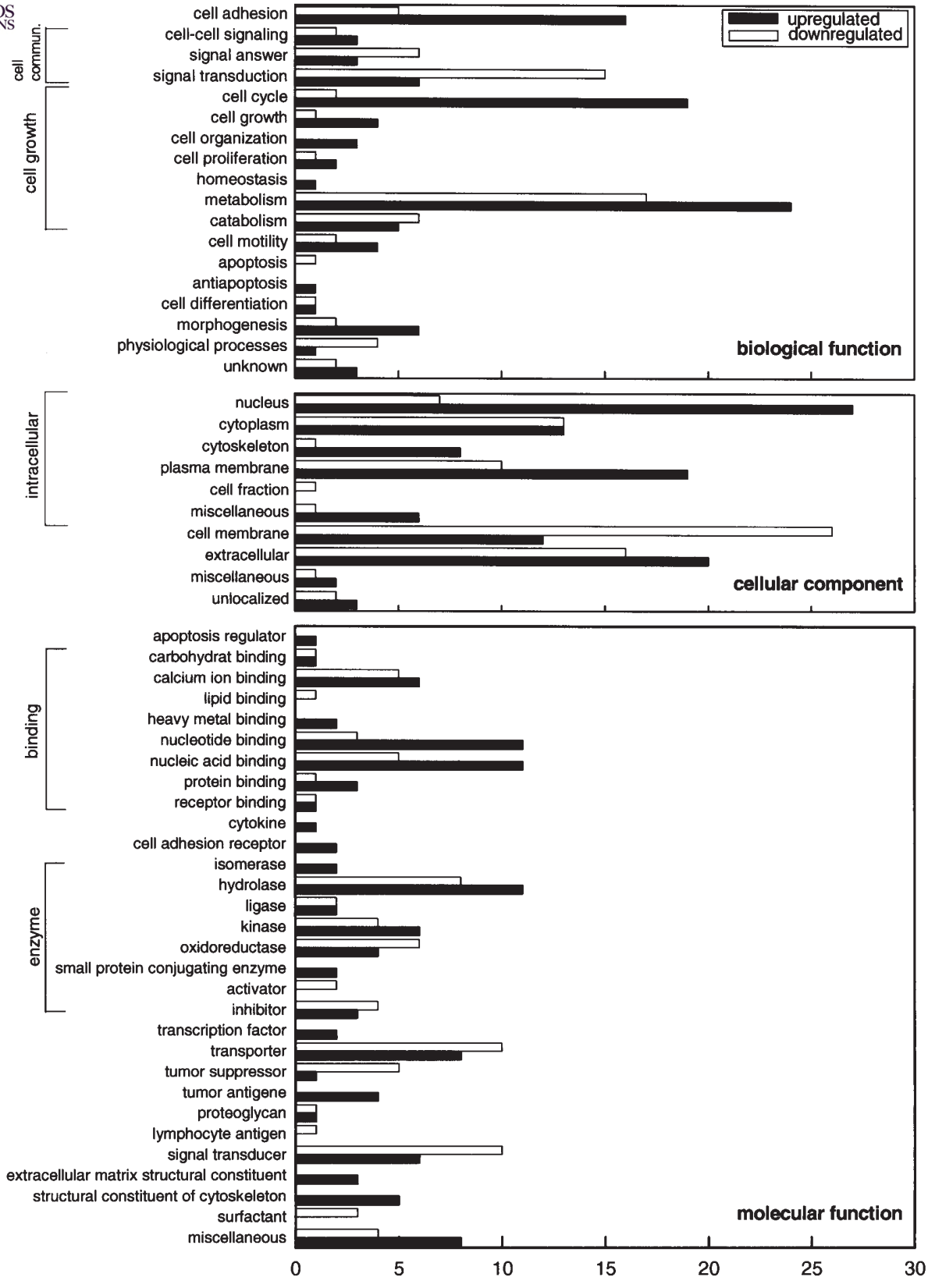



Figure 1. Biological function of differentially expressed genes in NSCLC specimens.

The Cox proportional hazards model was applied for multivariate analysis using the SPSS software program (SPSS Inc., Chicago, USA). Only patients with confirmed postoperative R0 status were admitted to the analysis of the disease-free survival. The threshold for statistical significance was p=0.05.

## Results

*Differentially expressed genes are related to specific biological functions*. Out of the roughly 59,000 probe sets on our oligonucleotide array 43,770 (73.5%) genes or ESTs showed

Table II. Significantly up-regulated genes in adenocarcinomas (ACC) and squamous cell carcinomas (SCC).

| Gene name | Expression AC | Expression SCC | P-value | Fold-change SCC/AC |
|---|---|---|---|---|
| Cytokeratin 6A (CK 6A) | 117.0 | 851.6 | $1.3x10^{-18}$ | 7.3 |
| Tumor protein 63 kDa | 71.4 | 421.9 | $3.0x10^{-17}$ | 5.9 |
| Keratin 16 | 79.2 | 458.5 | $1.9x10^{-8}$ | 5.8 |
| Desmocollin 3 | 60.5 | 233.2 | $1.4x10^{-11}$ | 3.9 |
| S 100 calcium-binding protein A2 | 155.6 | 588.1 | $2.9x10^{-14}$ | 3.8 |
| Cytokeratin 15 (CK 15) | 188.38 | 657.5 | $1.8x10^{-9}$ | 3.5 |
| Cytokeratin 17 (CK 17) | 68.6 | 238.2 | $4.9x10^{-12}$ | 3.5 |
| Collagen, Typ XVII/α 1 (COL17A1) | 106.4 | 302.1 | $4.2x10^{-7}$ | 2.8 |
| Collagen, Typ VII/α 1 (COL7A1) | 135.2 | 369.9 | $3.2x10^{-14}$ | 2.7 |
| FAT tumor suppressor homologue 2 (*Drosophila m.*) | 157.2 | 324.8 | $4.0x10^{-13}$ | 2.1 |
| Integrin α 6 | 109.6 | 224.1 | $6.0x10^{-9}$ | 2.0 |
| Thyroid transcription factor 1 | 379.0 | 94.6 | $6.0x10^{-11}$ | 0.25 |
| Specific transcription factor DAT1 | 182.3 | 64.5 | $6.4x10^{-8}$ | 0.35 |
| Paraoxonase 3 | 314.9 | 127.3 | $4.0x10^{-9}$ | 0.4 |
| Claudin 3 | 571.6 | 256.2 | $6.5x10^{-11}$ | 0.44 |
| Transcription factor 2/hepatic | 149.3 | 67.0 | $4.6x10^{-8}$ | 0.44 |
| Cytochrome P450 2B7 (CYP2B7) | 195.7 | 90.8 | $2.8x10^{-8}$ | 0.46 |
| Occludin | 201.2 | 96.7 | $1.8x10^{-8}$ | 0.48 |

an expression differences between NSCLC and normal lung over the dialed significance level and fold-change. At an approximate level of $p=8.0x10^{-7}$, we found only 187 (0.3%) genes, which were at least 2-fold up-regulated and 157 (0.35%) genes 2-fold down-regulated; 14,317 (24%) genes showed a signal intensity smaller than 60 in all samples, which was termed as absent calls.

Subsequently, all differentially regulated genes were classified into the biological pathway, cellular component and molecular function using the GO classification (Fig. 1). Among the genes with increased expression we determined cyclin-B1, cyclin-B2 and topoisomerase IIα, which are involved in cell cycle regulation. Transcripts encoding gene products associated with cell adhesion processes were also frequently overexpressed in lung tumors, such as collagens (COL1A1, COL3A1, COL7A1, COL11A1, COL17A1), desmocollins (desmocollin 2 and 3) and claudins (claudin 1 and 10). Additionally, we revealed up-regulated genes involved in catabolic processes of the extracellular matrix, including matrixmetalloproteinases (MMPs)-1, -9 and -12. Down-regulated genes are often related to signal transduction pathways (e.g., the receptor for advanced glycosylation end products (RAGE), and the complement component 5 receptor).

*Specific targets discriminate lung neoplasms*. Among the differently expressed genes we searched for stable targets that reliable discriminate the histological subtypes of NSCLC, adenocarcinoma and squamous cell carcinoma. At the defined significance level of $p=8.0x10^{-7}$, we found only 55 genes 2-fold up-regulated in squamous cell carcinoma in relation to the adenocarcinoma and 17 genes which were 2-fold up-regulated in adenocarcioma compared to squamous cell carcinoma (Table II). Squamous cell carcinomas were especially determined by elevated expressed genes encoding structural and filament proteins like cytokeratins (CK 6A, 15, 17), collagens (COL7A1, COL17A1) and adhesions proteins (desmocollin 3, integrin α 6). Also, carcinogenesis-related factors like the tumor protein p63 and FAT tumor suppressor homolog 2 (*Drosophila melanogaster*) were more than 2-fold elevated in squamous cell carcinomas.

Among the small group of up-regulated genes in adenocarcinomas three transcription factors were identified: thyroid transcription factor 1 (TTF1), specific transcription factor DAT1 and transcription factor 2/hepatic. Claudin 3, which encodes an adhesion protein, was also significantly increased in adenocarcinomas. However, based on the distinct heterogeneity of tumour cells/cellular heterogeneity in lung tumours all these 'specific' genes had a high coexpression in the complementary histological subtype. TTF1 was also strongly expressed in 24.5% of the squamous cell carcinomas. Thus, we calculated the expression ratio per sample by dividing the expression value of TTF1 and CK 6A. The dispersion of values after ratio calculation remarkably decreased (Fig. 2). On the basis of this ratio technique, we accurately discriminated 95% of blinded tumor probes (n=83) as squamous cell carcinoma or adenocarcinoma (data not shown).

*Expressions of MMP-12 and uPA correlated with relapse-free survival and survival*. Proteolysis plays an important role in the multiple-step process of metastasis. Most of the investigated genes associated with proteolysis, e.g. serine-, cysteinyl-

SPANDIDOS PUBLICATIONS

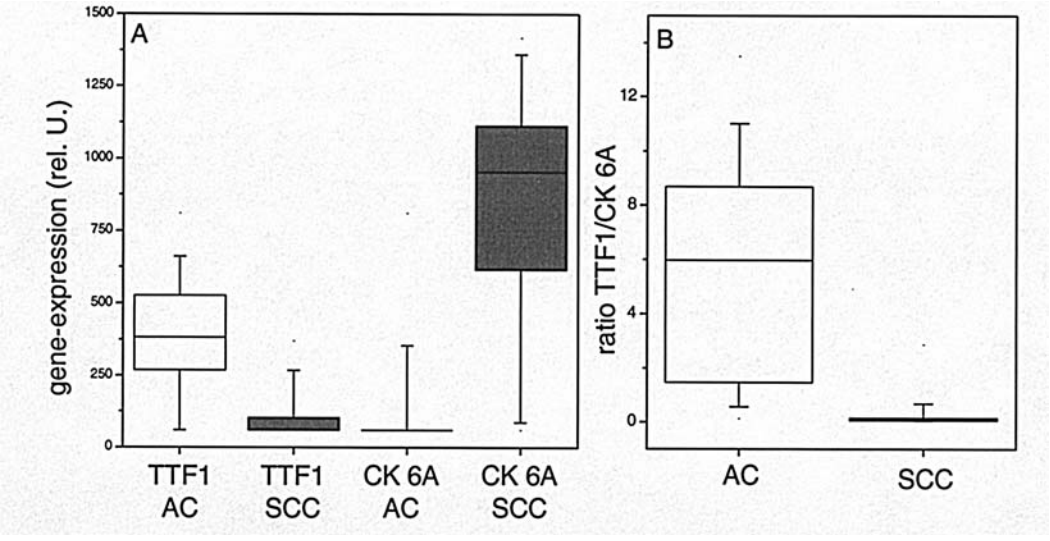



Figure 2. Box-and-whisker plots of gene expression of thyroid transcription factor 1 (TTF1) and cytokeratin 6A (CK 6A) (A) and the ratio of TTF1/CK 6A (B) for adenocarcinomas (AC) and squamous cell carcinomas (SCC).

Table III. Gene expression data of selected genes involved in metastases.

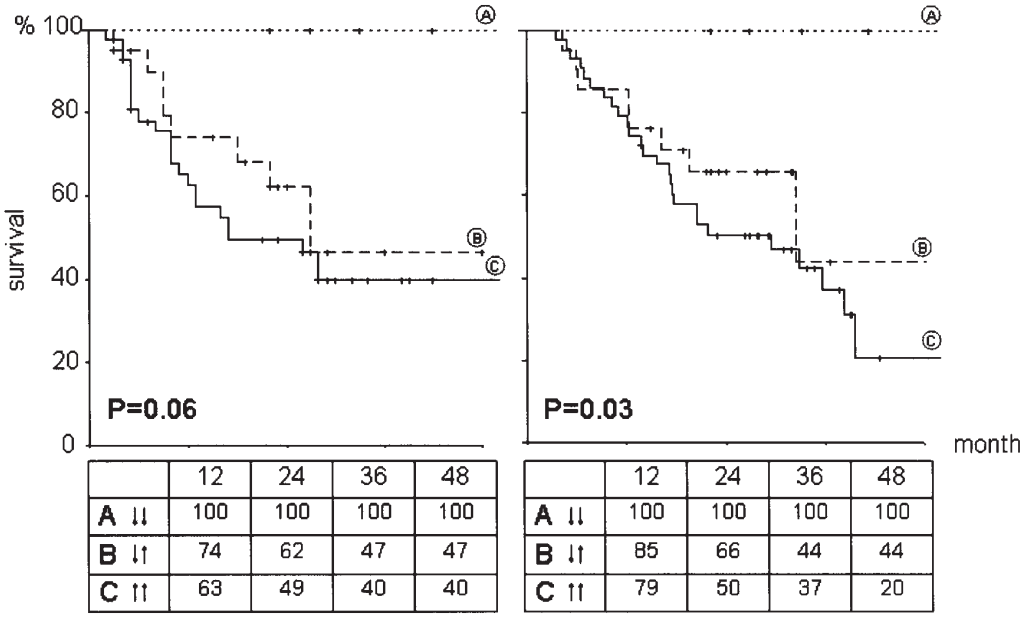
| | Expression lung | Expression tumor | *t*-test | Fold-change tumor/lung |
|---|---|---|---|---|
| Matrix metalloproteinases | | | | |
| MMP-1 | 60.0 | 220.0 | $1.9x10^{-11}$ | 3.7 |
| MMP-9 | 252.0 | 530.0 | $5.5x10^{-19}$ | 2.1 |
| MMP-10 | 60.0 | 185.6 | $3.1x10^{-6}$ | 3.0 |
| MMP-11 | 213.9 | 342.8 | $7.9x10^{-8}$ | 1.6 |
| MMP-12 | 60.0 | 248.7 | $1.4x10^{-20}$ | 4.1 |
| Serine proteinases | | | | |
| uPA-α | 307.0 | 493.0 | $1.9x10^{-10}$ | 1.6 |
| Cysteine proteinases | | | | |
| Cathepsin B | 1044.5 | 1086.0 | 0.39 | 1.0 |
| Cathepsin L | 707.0 | 618.0 | $5.0x10^{-4}$ | 0.87 |
| Aspartic proteinases | | | | |
| Cathepsin D | 78.5 | 78.0 | 0.89 | 1.0 |

and aspartyl-proteinases as well as matrix metalloproteinases (MMPs), remained unaltered in NSCLC samples (Table III), however, MMP-1, MMP-9 and MMP-12 were increased at significant level. MMP-10, MMP-11 and urokinase plasminogen activator α (uPA) expression achieved only one of the quality objectives for differentially expressed genes, level of significance or fold-change. Because the expression values of these genes were highly variable in the tumor samples, we individually determined cut-offs for the classification into groups of low and high expression. The cut-offs were defined by an expression level of about 150% of the control lung samples (cut-off for MMP-1, -10, -12 set to 100, for MMP-11 set to 300 and for MMP-9 and uPA set to 400 relative expression units, respectively). Subsequently, these cut-off values were used for detailed clinical evaluation of the gene array data. We revealed that 29 (51%) out of the 56 patients with high expression of MMP-12 developed local recurrence (n=4), distant metastasis (n=19) or both (n=6). The tumor relapse rate in the group of low MMP-12 expression was only 14% (n=2); one patient had local recurrence and one patient local recurrence and distant metastasis. Patients with low MMP-12 showed a significant longer median tumor relapse-free survival time (p=0.02) as well as a longer total median survival after operation (p=0.01) in relation to patients with MMP-12 high expressed tumors (Table IV).

Lung tumor samples of patients with high uPA expression also tend to have higher rates of tumor relapse. Nearly 50% (25 of 51) of patients with high expression of uPA had local

Table IV. MMP-12 and uPA in correlation to tumor relapse-free survival and total survival.

| | Low expression | High expression | P-value |
|---|---|---|---|
| MMP-12 | | | |
| N | 14 | 56 | |
| Mean tumor relapse-free survival (months) | 48.6 | 29.1 | |
| 4-year tumor relapse-free survival (%) | 85.1 | 39.8 | 0.02 |
| Mean survival time (months) | 48.8 | 28.6 | |
| 4-year survival rate (%) | 85.7 | 20.8 | 0.01 |
| uPA-α | | | |
| N | 19 | 51 | |
| Mean tumor relapse-free survival (months) | 40.9 | 29.7 | |
| 4-year tumor relapse-free survival (%) | 61.5 | 43.6 | 0.1 |
| Mean survival time (months) | 41.8 | 29.5 | |
| 4-year survival rate (%) | 61.0 | 26.4 | 0.09 |




| | 12 | 24 | 36 | 48 |
|---|---|---|---|---|
| A ↓↓ | 100 | 100 | 100 | 100 |
| B ↓↑ | 74 | 62 | 47 | 47 |
| C ↑↑ | 63 | 49 | 40 | 40 |

| | 12 | 24 | 36 | 48 |
|---|---|---|---|---|
| A ↓↓ | 100 | 100 | 100 | 100 |
| B ↓↑ | 85 | 66 | 44 | 44 |
| C ↑↑ | 79 | 50 | 37 | 20 |

Figure 3. Kaplan-Meier plots of disease-free survival (left) for R0 resected NSCLC patients and total survival (right); group A (n=6), low expression of MMP-12 and uPA; group B (n=21), low expression of only one gene (MMP-12 or uPA) and high expression of the other; group C (n=43), high expression of MMP-12 and uPA.

recurrence and/or distant metastasis, whereas only 31% of patients (6/19) with low expression of uPA developed tumor relapse (p=0.2). Kaplan-Meier analysis of the tumor relapse-free survival and total survival revealed differences for uPA between low and high expressed tumors (Table IV). Although there was no correlation between the expression of MMP-12 and uPA we created three groups of patients depending on the combined gene expression of MMP-12 and uPA. This characterization revealed that all patients with low expression of both genes (group A) survived without tumor relapse for 4 years, whereas patients with high expression of these two genes (group C) had the lowest tumor-free survival rate as well as the lowest 4-year survival rate (20%). Patients with high expression of only one of these two genes (group B) had a tumor-free survival rate of 47% and a total survival rate of 44%. The differences between these three groups were significant (p=0.03) for the total survival (Fig. 3).

## Discussion

The DNA microarray technology provides the technical conditions to obtain a genome-wide picture of cells and tissues. Microarray-based determination of gene expression profiles may, on the one hand, identify new malignancy-relevant

SPANDIDOS PUBLICATIONS

d signal pathways; on the other hand, it offers the ity to establish new tumor classifications and perhaps molecular staging systems (15-17). Several gene array studies of the major lung cancer types have been published including data of antibody microarrays (18). However, up to now these studies worked with global DNA-arrays and no clinical relevant changes followed. The objective of the present work was to evaluate differentially expressed genes as potential targets for a lung cancer specific DNA-array for new diagnosis and therapy concepts.

Out of 59,620 examined genes or ESTs, only 0.6% (n=344) were expressed significantly differentially between lung and tumor, which could be an indication that only few genes change their expression during the processes of tumorigenesis and metastases. In the literature the proportions of differentially regulated genes are specified to different degrees ranging from 0.08% in prostate carcinoma (19) to 20% in breast cancer (20). In this regard, the results are largely influenced by the selection of the genes located on the DNA chip. So, tumor-specific gene chips (21) show more expression differences of their spotted sequences than gene chips, representing, e.g. a large part of the human genome.

Besides the number of the differentially expressed genes, the function of the gene products is essential. The classification according to biological process gives insights into molecular changes occurring in tumor development and progression. In the present work, the largest part of the highly expressed genes in the tumor tissues was involved in the processes of cell growth, including metabolism and the cell cycle, and in the processes of cell communication and adhesion.

Ongoing activation of cell cycle/proliferation represents one of the decisive features of tumor development. Therefore, genes involved in cell cycle regulation showed a higher expression in almost all DNA microarray analyses studying bronchial carcinomas. Cyclines (including cyclin-B2) and cycline-dependent kinases are very often overexpressed in lung cancer (22-26). The highly expressed cyclines-B1 and -B2 play a key role in the G2/M phase. It has been reported that patients whose tumor expressed cyclin-B1 at a high level had a significantly shorter survival time than did patients whose tumors expressed cyclin-B1 at a low level (27).

DNA topoisomerases including TOP2A are repeatedly shown as highly expressed genes in NSCLC (22) and SCLC (23). Chen *et al* were able to prove that an increasing metastasising capacity of pulmonary adenocarcinoma cell lines is associated with an increased expression of TOP2A (28). On the basis of topoisomerase inhibitors as clinically applied anticancer drugs, the detection of DNA topoisomerases may also be useful as the therapeutic target (29). Moreover, it was possible to show *in vitro* that the sensitivity towards topoisomerase inhibitors correlate with the expression level of topoisomerase II α in malignant cells (30).

Genes involved in cell adhesion were also up-regulated in the tumor samples. The highly expressed collagens (COL1A1, COL3A1 and COL11A1) are relatively ubiquitously represented in all NSCLC and might reflect the stroma proportion in the tumor tissue. Nevertheless, other collagens (COL7A1 and COL17A1) are specific for histological subtypes (e.g. squamous cell carcinoma) of the NSCLC. The increased expression of COL7A1 and COL17A1 seems to be a sign of squamous cell differentiation of the tumor, as they were also found in carcinomas of the head-and-neck area, the vulva, vagina (31) and the skin, the oesophagus and cervix (32). Out of other highly expressed genes coding for cell adhesion molecules, Claudin 10 and DSC 2 and 3 seem to play a significant role in NSCLC. The gene product of Claudin 10 is secreted and might therefore be used as a potential serum tumor marker (33). In contrast to Claudin, the expression of DSC2 and DSC3 depends on histological subtypes. While adenocarcinomas only express DSC2, squamous cell carcinomas are characterized by an increased gene expression of DSC2 and DSC3 (34).

Gene expression differences between normal and tumor tissues including histological subtypes provide the opportunity of automated tissue discrimination with the aim of automated diagnostics. The present study proved that the combination of gene pairs, by applying expression ratios and cut-offs, does provide significantly more information than the expression of an individual gene. This has also been described for tissue differentiation of tumor and normal lung by calculating the expression ratio of RAGE and cyclin-B2 (35). We have additionally shown the discrimination of squamous cell carcinomas and adenocarcinomas by determing the ratio of TTF1 per CK-6A. This expression ratio technique represents a more efficient method to translate the strengths of microarray expression profiling into easy clinical tests. Moreover, this technique is simple with broad clinical use in diagnosis as well as prediction of prognosis in cancer. The accuracy is even higher if the expression value of the two differentially expressed genes is nearly the same. Additionally, this ratio-based technique is independent from the expression measuring method, needs no gene expression reference (house-keeping gene as loading control), requires only a small amount of RNA and does not require the coupling of transcription to translation for the chosen genes (36). Gordon *et al* tested expression ratios in the discrimination of two different tissues for the first time (37). Using two or three expression ratios of two differentially expressed genes, they found that the differential diagnosis of mesothelioma and pulmonary adenocarcinoma was 95 and 99% accurate, respectively. With the geometric mean of three most accurate individual ratios of four genes Bueno *et al* were able to distinguish normal prostate and prostate cancer samples obtained at surgery with 90% accuracy (38). This technique is highly precise in the discrimination of cancer tissues, and it can be equally useful in additional clinical applications such as prediction of outcome. In mesothelioma patients treatment-related outcome was predicted by gene expression ratio-based analysis (36).

So far the expression changes of many individual genes were possible to be correlated with the prognosis of tumor patients (39). However, in most of the cases the reproducibility of the results in larger patient groups was missing, as a single prognosis factor may not reflect the complexity of tumorigenesis. As a conclusion of these findings, individual predictions of a possible relapse and/or a survival prognosis are probably only possible by using a group of genes. In the present work R0-resected patients could be divided in groups

with low to high metastases formation risk based on the expression of two genes (MMP-12, uPA). It was possible to show that all R0 resected patients with tumors having low levels of MMP-12 and uPA did not incur any risk of a relapse in the first 4 years after the operation. MMP-12 and uPA in combination also allowed for a statement on total survival. Only one report exists so far on the NSCLC, in which two factors, MRP-1/CD9 gene expression and the K-ras gene mutation, are applied simultaneously to determine the survival prognosis (40).

The high expression of MMPs and uPA provides the opportunity of using these genes for target-based immunotherapies. First generations of MMP inhibitors (MMPIs) already proved an extensive broad-spectrum MMP blockade (e.g. batimastat, marimastat, prinomastat). In contrast, MMPIs of the last generations resulted in a more selective blockade of distinct MMPs. BAY 12-9566 selectively inhibits MMP-2, -3 and -9 while ONO-4817 impairs the function of MMP-2, -8, -9, -12 and -13 by blocking the zinc binding sites. The application of anti-uPA antibodies considerably reduced the matrix invasion and the formation of lung metastases in a Lewis Lung cancer model (41). Several other studies proved an *in vivo* anti-tumor effect through treatment with amilorid (42,43), which is a non-specific uPA inhibitor.

In the clinical practice, the most commonly used target genes for treating NSCLCs are EGFR and VEGF (44). Both target genes are affected by immunotherapeutics in combination with classic chemotherapeutics (Sandler AB, *et al*, ASCO, Abst, 2005). Nonetheless, the clinical application of those immunotherapeutics and other targeted treatments of cancer have so far been disappointing. This may due to the non-specificity of the inhibitory agents and the lack of relevance of these drugs to patients without individual target induction. Therefore, diagnostics and treatment strategies should include the results of tumor-specific DNA chips, which only analyze a low but relevant number of specific target genes. Measuring the genetic profile of the respective tumors may consequently lead to improvements in diagnostics, therapy decisions, and success of anticancer therapies.

## Acknowledgements

This study was supported by grants from EOS Biotechnology, DFG SFB 610 TPB1, SFB 598 TP A5 and Deutsche Krebshilfe (70-2787-Bu 3) and BMBF (01-ZZ0109). The sponsors of the study had no role in data interpretation or writing of the report.

## References

1. Jemal A, Thomas A, Murray T and Thun M: Cancer statistics, 2002. CA Cancer J Clin 52: 23-47, 2002.
2. Steels E, Paesmans M, Berghmans T, Branle F, Lemaitre F, Mascaux C, Meert AP, Vallot F, Lafitte JJ and Sculier JP: Role of p53 as a prognostic factor for survival in lung cancer: a systematic review of the literature with a meta-analysis. Eur Respir J 18: 705-719, 2001.
3. Han H, Landreneau RJ, Santucci TS, Tung MY, Macherey RS, Shackney SE, Sturgis CD, Raab SS and Silverman JF: Prognostic value of immunohistochemical expressions of p53, HER-2/neu and bcl-2 in stage I non-small-cell lung cancer. Hum Pathol 33: 105-110, 2002.
4. Liao M, Wang H, Lin Z, Feng J and Zhu D: Vascular endothelial growth factor and other biological predictors related to the postoperative survival rate on non-small cell lung cancer. Lung Cancer 33: 125-132, 2001.
5. Shiba M, Kohno H, Kakizawa K, Iizasa T, Otsuji M, Saitoh Y, Hiroshima K, Ohwada H and Fujisawa T: Ki-67 immunostaining and other prognostic factors including tobacco smoking in patients with resected non-small cell lung carcinoma. Cancer 89: 1457-1465, 2000.
6. Martin B, Paesmans M, Berghmans T, Branle F, Ghisdal L, Mascaux C, Meert AP, Steels E, Vallot F, Verdebout JM, Lafitte JJ and Sculier JP: Role of Bcl-2 as a prognostic factor for survival in lung cancer: a systematic review of the literature with meta-analysis. Br J Cancer 89: 55-64, 2003.
7. Staege MS, Hattenhorst UE, Neumann UE, Hutter C, Foja S and Burdach S: DNA-microarrays as tools for the identification of tumor specific gene expression profiles: applications in tumor biology, diagnosis and therapy. Klin Padiatr 215: 135-139, 2003.
8. Travis W, Linder J and Mackay B: Classification, histology, cytology and electon microscopy. In: Lung Cancer: Principles and Practice. Pass H, Mitchel J, Johnson D and Turisi A (eds). Lippincott-Raven, Philadelphia, pp361-395, 1996.
9. Mountain C: Revision in the international system for staging lung cancer. Chest 111: 1170-1117, 1997.
10. Wodicka L, Dong H, Mittmann M, Ho MH and Lockhart DJ: Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat Biotechnol 15: 1359-1367, 1997.
11. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531-537, 1999.
12. Lipshutz RJ, Fodor SP, Gingeras TR and Lockhart DJ: High density synthetic oligonucleotide arrays. Nat Genet 21: 20-24, 1999.
13. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA and Wheeler DL: GenBank. Nucleic Acids Res 28: 15-18, 2000.
14. Salamov AA and Solovyev VV: Ab initio gene finding in *Drosophila* genomic DNA. Genome Res 10: 516-522, 2000.
15. Ludwig JA and Weinstein JN: Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer 5: 845-856, 2005.
16. Staege MS, Hutter C, Neumann I, Foja S, Hattenhorst UE, Hansen G, Afar D and Burdach SE: DNA microarrays reveal relationship of Ewing family tumors to both endothelial and fetal neural crest-derived cells and define novel targets. Cancer Res 64: 8213-8221, 2004.
17. Staege MS, Hansen G, Baersch G and Burdach S: Functional and molecular characterization of interleukin-2 transgenic Ewing tumor cells for *in vivo* immunotherapy. Pediatr Blood Cancer 43: 23-34, 2004.
18. Bartling B, Hofmann HS, Boettger T, Hansen G, Burdach S, Silber RE and Simm A: Comparative application of antibody and gene array for expression profiling in human squamous cell lung carcinoma. Lung Cancer 49: 145-154, 2005.
19. Magee JA, Araki T, Patil S, Ehrig T, True L, Humphrey PA, Catalona WJ, Watson MA and Milbrandt J: Expression profiling reveals hepsin overexpression in prostate cancer. Cancer Res 61: 5692-5696, 2001.
20. Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415: 530-536, 2002.
21. Inoue H, Matsuyama A, Mimori K, Ueo H and Mori M: Prognostic score of gastric cancer determined by cDNA microarray. Clin Cancer Res 8: 3475-3479, 2002.
22. Heighway J, Knapp T, Boyce L, Brennand S, Field JK, Betticher DC, Ratschiller D, Gugger M, Donovan M, Lasek A and Rickert P: Expression profiling of primary non-small cell lung cancer for target identification. Oncogene 21: 7749-7763, 2002.
23. Bangur CS, Switzer A, Fan L, Marton MJ, Meyer MR and Wang T: Identification of genes overexpressed in small cell lung carcinoma using suppression subtractive hybridization and cDNA microarray expression analysis. Oncogene 21: 3814-3825, 2002.

SPANDIDOS PUBLICATIONS

Kapoor M, Newton K, Cheon K, Ramaswamy A, Lotan R, reveals LC and Koo JS: Global detection of molecular changes concurrent alteration of several biological pathways in non-small cell lung cancer cells. Mol Genet Genomics: 1-14, 2005.
25. Kettunen E, Anttila S, Seppanen JK, Karjalainen A, Edgren H, Lindstrom I, Salovaara R, Nissen AM, Salo J, Mattson K, Hollmen J, Knuutila S and Wikman H: Differentially expressed genes in non-small cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer. Cancer Genet Cytogenet 149: 98-106, 2004.
26. Amatschek S, Koenig U, Auer H, Steinlein P, Pacher M, Gruenfelder A, Dekan G, Vogl S, Kubista E, Heider KH, Stratowa C, Schreiber M and Sommergruber W: Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes. Cancer Res 64: 844-856, 2004.
27. Soria JC, Jang SJ, Khuri FR, Hassan K, Liu D, Hong WK and Mao L: Overexpression of cyclin B1 in early-stage non-small cell lung cancer and its clinical implication. Cancer Res 60: 4000-4004, 2000.
28. Chen JJ, Peck K, Hong TM, Yang SC, Sher YP, Shih JY, Wu R, Cheng JL, Roffler SR, Wu CW and Yang PC: Global analysis of gene expression in invasion by a lung cancer model. Cancer Res 61: 5223-5230, 2001.
29. Topcu Z: DNA topoisomerases as targets for anticancer drugs. J Clin Pharm Ther 26: 405-416, 2001.
30. Isaacs RJ, Davies SL, Sandri MI, Redwood C, Wells NJ and Hickson ID: Physiological regulation of eukaryotic topoisomerase II. Biochim Biophys Acta 1400: 121-137, 1998.
31. Wetzels RH, Schaafsma HE, Leigh IM, Lane EB, Troyanovsky SM, Wagenaar SS, Vooijs GP and Ramaekers FC: Laminin and type VII collagen distribution in different types of human lung carcinoma: correlation with expression of keratins 14, 16, 17 and 18. Histopathology 20: 295-303, 1992.
32. Yamada T, Endo R, Tsukagoshi K, Fujita S, Honda K, Kinoshita M, Hasebe T and Hirohashi S: Abberant expression of a hemidesmosomal protein, bullous pemphigoid antigen, in human squamous cell carcinoma. Lab Invest 75: 589-600, 1996.
33. Sugita M, Geraci M, Gao B, Powell RL, Hirsch FR, Johnson G, Lapadat R, Gabrielson E, Bremnes R, Bunn PA and Franklin WA: Combined use of oligonucleotide and tissue microarrays identifies cancer/testis antigens as biomarkers in lung carcinoma. Cancer Res 62: 3971-3979, 2002.
34. Nuber UAL: Expression der Gene der Desmocollin-Familie in menschlichen Zellen und Geweben. Dissertation A, Medizinsche Fakultät, Heidelberg, 1999.
35. Hofmann HS, Hansen G, Burdach S, Bartling B, Silber RE and Simm A: Discrimination of human lung neoplasm from normal lung by two target genes. Am J Respir Crit Care Med 170: 516-519, 2004.
36. Gordon GJ, Jensen RV, Hsiao LL, Gullans SR, Blumenstock JE, Richards WG, Jaklitsch MT, Sugarbaker DJ and Bueno R: Using gene expression ratios to predict outcome among patients with mesothelioma. J Natl Cancer Inst 95: 598-605, 2003.
37. Gordon GJ, Jensen RV, Hsiao LL, Gullans SR, Blumenstock JE, Ramaswamy S, Richards WG, Sugarbaker DJ and Bueno R: Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. Cancer Res 62: 4963-4967, 2002.
38. Bueno R, Loughlin KR, Powell MH and Gordon GJ: A diagnostic test for prostate cancer from gene expression profiling data. J Urol 171: 903-906, 2004.
39. Hofmann HS, Hansen G, Richter G, Taege C, Simm A, Silber RE and Burdach S: Matrix metalloproteinase-12 expression correlates with local recurrence and metastatic disease in non-small cell lung cancer patients. Clin Cancer Res 11: 1086-1092, 2005.
40. Miyake M, Adachi M, Huang C, Higashiyama M, Kodama K and Taki T: A novel molecular staging protocol for non-small cell lung cancer. Oncogene 18: 2397-2404, 1999.
41. Kobayashi H, Gotoh J, Shinohara H, Moniwa N and Terao T: Inhibition of the metastasis of Lewis lung carcinoma by antibody against urokinase-type plasminogen activator in the experimental and spontaneous metastasis model. Thromb Haemost 71: 474-480, 1994.
42. Jankun J, Keck RW, Skrzypczak-Jankun E and Swiercz R: Inhibitors of urokinase reduce size of prostate cancer xenografts in severe combined immunodeficient mice. Cancer Res 57: 559-563, 1997.
43. Evans DM, Sloan-Stakleff K, Arvan M and Guyton DP: Time and dose dependency of the suppression of pulmonary metastases of rat mammary cancer by amiloride. Clin Exp Metastasis 16: 353-357, 1998.
44. Hofmann HS and Bartling B: How will lung cancer be treated in the future? Future Oncol 1: 551-559, 2005.