Molecular profiles of non-small cell lung cancers in cigarette smoking and never-smoking patients

Szymanowska-Narloch A¹, Jassem E¹, Skrzypski M², Muley T³, Meister M³, Dienemann H³, Taron M⁴, Rosell R⁴, Rzepko R⁵, Jarząb M⁶, Marjański T⁷, Pawłowski R⁸, Rzyman W⁷, Jassem J²,*

¹ Department of Allergology, Medical University of Gdansk, Gdansk, Poland
² Department of Oncology and Radiotherapy, Medical University of Gdansk, Gdansk, Poland
³ Thoraxklinik, University Hospital Heidelberg, Translational Research Center Heidelberg, Heidelberg, Germany
⁴ Catalan Institute of Oncology, Hospital Germans Trias i Pujol, Barcelona, Spain
⁵ Department of Pathology, Medical University of Gdansk, Gdansk, Poland
⁶ Radiotherapy and Chemotherapy Clinic, Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, Gliwice, Poland
⁷ Department of Thoracic Surgery, Medical University of Gdansk, Gdansk, Poland
⁸ Department of Forensic Medicine, Medical University of Gdansk, Gdansk, Poland

* CORRESPONDING AUTHOR:
Department of Oncology and Radiotherapy
Medical University of Gdansk,
Dębinki 7,
80-211 Gdansk, Poland
Tel.: +4858 349 2270
Fax: +4858 349 2210
E-mail: jjassem@umURED.edu.pl (Jacek Jassem)

ABSTRACT

Purpose: Molecular features of non-small cell lung cancer (NSCLC) in never-smokers are not well recognized. We assessed the expression of genes potentially related to lung cancer etiology in smoking vs. never-smoking NSCLC patients.

Methods: We assayed frozen tumor samples from surgically resected 31 never-smoking and 54 clinically pair-matched smoking NSCLC patients, and from corresponding normal lung tissue from 27 and 43 patients, respectively. Expression of 21 genes, including cell membrane kinases, sex hormone receptors, transcription factors, growth factors and others was assessed by reverse transcription – quantitative PCR.

Results: Expression of 5 genes was significantly higher in tumors of non-smokers vs. smokers: CSF1R (p<0.0001), RRAD (p<0.0001), PR (p=0.0004), TGFBR2 (p=0.0027) and EPHB6 (p=0.0033). Expression of AKR1B10 (p<0.0001), CDKN2A (p<0.0001), CHRNA6 (p<0.0001), SOX9 (p<0.0001), survivin (p<0.0001) and ER2 (p=0.002) was significantly higher in tumors compared to normal lung tissue. Expression of AR (p<0.0001), EPHB6 (p<0.0001), PR (p<0.0001), TGFBR2 (p<0.0001), TGFBR3 (p<0.0001), ER1 (p=0.0006) and DLG1 (p=0.0016) was significantly lower in tumors than in normal lung tissue. Expression of IGF2 was higher in tumors than in healthy lung tissue in never-smokers (p<0.0001), and expression of AHR (p<0.0001), CSF1R (p<0.0001) and RRAD (p<0.0001) was lower in tumors than in healthy lung tissue in smokers.

Conclusion: Expression of several genes in NSCLC is strongly related to smoking history. Lower expression of PR and higher expression of ER2 in tumors suggests a possibility of hormonal therapeutic intervention in selected NSCLC patients. Distinct molecular features of NSCLC in never-smokers, e.g. CHRNA6 upregulation, may prompt new treatment strategies.

Key words: Non-small cell lung cancer, gene expression, smoking

INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality [1]. Tobacco smoke is a well-recognized cause of non-small cell lung cancer (NSCLC), however about 15% of men and 50% of women diagnosed with lung cancer worldwide have never smoked [2]. The tumors of non-smokers are predominantly located in the peripheral lung areas and are more likely to originate from bronchoalveolar stem cells [3].
The never-smokers with lung cancer are more frequently women than men.

NSCLCs of non-smokers are also distinct in respect to molecular pathology, e.g. they are more likely to harbor EGFR activating mutations [4] and a fusion oncogene EMLA-ALK [5]. There are also differences in gene expression in NSCLCs and in normal lung tissue of smokers and never-smokers [6,7]. Further molecular characterization of NSCLC in never-smoking patients might aid the development of new therapeutic strategies.

In this study, we assessed the expression of selected genes potentially driving tumorigenesis and biology of NSCLC in tumors and normal lung tissue of never-smoking patients, using reverse transcription – quantitative polymerase chain reaction (RT-qPCR). The genes were selected based on a thorough current literature search. Hormone receptor genes were investigated, as the use of hormone replacement therapy (HRT) in women was shown to be associated with an earlier onset of NSCLC and shorter overall survival (OS) [8]. Another study indicated that combined analysis of estrogen receptor beta (ERβ1) and progesterone receptor (PR) expression identified lung cancer patients with poor outcome [9]. Thus, the sex hormones emerge not only as factors triggering lung oncogenesis but they may also impact the course of the disease. Lung is also a target tissue for androgen receptor (AR) which is expressed in different lung cancer types [10]. Given the above, we hypothesized that the sex hormones background might elicit the NSCLC in never-smokers. To verify such correlation, we analyzed the expression of major sex hormone receptors (ER1, ER2, PR and AR). Carcinogens in tobacco smoke were found to interact with estrogen and androgen receptors and to affect hormonal proliferative signaling balance [11]. Previous studies demonstrated that aryl hydrocarbon receptor (AHR) is overexpressed in lung adenocarcinomas [12]. Therefore, AHR, a transcription factor that is activated by cigarette smoke was included into the analysis. AKR1B10, CHRNA6, CHRNβ3 and CDKN2A/P16 genes were included, as their expression is affected by the tobacco smoke [13-17]. AKR1B10 expression is upregulated by smoking in airway epithelia in healthy smokers and this event might constitute an early step in lung carcinogenesis [13]. In a previous study by Lam et al. [14], the expression of CHRNA6 and CHRNβ3 were in normal lung cancer cells from non-smokers compared to smokers. CDKN2A/P16 is a tumor suppressor gene which is inactivated in many tumors including lung cancer [16]. On the other hand, Zhao et al. [17] demonstrated that in one third of immunohistochemistry P16-positive NSCLC cases, expression of this protein was caused by increased gene copy number of CDKN2A. This diverse and apparently molecular context-dependent expression of P16 in NSCLC prompted the inclusion of this transcript into the pool of analyzed genes. Crizotinib, EMLA-ALK fusion gene inhibitor, was recently shown to be efficacious in metastatic NSCLC harboring this abnormality, which is typical for never-smokers [18]. Oncogenic potential of ALK kinase may be acquired through gene fusion and possibly by gene amplification [19]. Thus, the transcript for ALK (covering all potential variants of oncogenic fusion transcripts containing ALK) [20] was assessed, together with 3 other factors of insulin receptor superfamily (IGF1, IGF1R and IGF2), potentially relevant to biology of NSCLC of non-smoking patients [21-23]. Finally, a number of genes (survivin, CSF1R, DLG1, EPHB6, RRAD, SOX9, TGFBR2, TGFBR3) were included in the analysis to validate previous findings [7, 24-30].

Due to a large size of the full article its Supplementary Files: Figures 1S – 8S (Appendix A) and Tables 1S–3S (Appendix B), are available only online at: http://dx.doi.org/10.2478/ams-2013-0025.

MATERIALS AND METHODS

Study subjects

This study was approved by the institutional review boards of all participating institutions. All patients underwent between 2002 and 2008 pulmonary resection with curative intent in two institutions (Department of Thoracic Surgery, Medical University of Gdansk, Poland, and Department of Surgery, Thoraxklinik at University Hospital Heidelberg, Germany), a member of the accredited National Center of Tumor Diseases tissue bank. Patients for whom snap-frozen lung tumor was available were enrolled in the study; for most of these patients also corresponding healthy lung tissue samples were available. None of the patients was administered preoperative or postoperative chemotherapy or radiotherapy (Tab. 1). Tumor staging was performed using AJCC/UICC classification from 1997. By the time of this analysis (February 2012), 42 out of 85 patients deceased.

The non-smoking group included 31 patients with pathologically confirmed NSCLC. Patients were defined as never-smokers if they declared smoking fewer than 100 cigarettes in their life-time. Each never-smoker was pair-matched to 1 or 2 smoking patients with similar clinical characteristics (age, sex, histological tumor type, pTNM). The reference group included 54 NSCLC patients with smoking history. Most of these patients had a history of more than 20 pack-years, with mean pack-years of 47.9.

Methods

Using RT-qPCR we compared expression of 21 genes (Tab. 2) in NSCLCs of never-smokers vs. smokers, and in tumors vs. corresponding normal lung tissue. These included genes associated with smoking (e.g. RRAD, CHRNA6, CHRNβ3, AHR, P16/CDKN2A), kinase receptors (CSF1R, EPHB6, ALK), hormonal receptors (e.g. PR, AR, ER1-2), growth factor receptors (e.g. TGFBR2, TGFBR3, IGF1R), apoptosis inhibitors (survivin) and others. Normalization genes
Sequence Detection System (Applied Biosystems). Each Density Arrays was performed using ABI PRISM 7900 HT activity and specific primers. Analysis with TaqMan Low (MFC), TaqMan probes, polymerase with 5'-exonuclease RT-qPCR was performed with the use of microfluidic cards cDNA. Quantitative analysis of mRNA of selected genes by hexameric primers. One μg of RNA was transcribed into cDNA Archive Kit, Applied Biosystems) with the use of chain reaction rewriting RNA to cDNA (High Capacity –CHRNA).

Reverse transcription was combined with the polymerase chain reaction rewriting RNA to cDNA (High Capacity cDNA Archive Kit, Applied Biosystems) with the use of hexameric primers. One μg of RNA was transcribed into cDNA. Quantitative analysis of mRNA of selected genes by RT-qPCR was performed with the use of microfluidic cards (MFC), TaqMan probes, polymerase with 5′-exonuclease activity and specific primers. Analysis with TaqMan Low Density Arrays was performed using ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems). Each channel of a microfluidic card was loaded with a 100 μl mix of equal volume of Taqman Universal PCR Master Mix (Applied Biosystems) and cDNA template (corresponding to 100 ng of total RNA).

### Statistical analysis

RT-qPCR results were obtained using the Sequence Detection System 2.1 Software. Gene expression was calculated with 2ΔΔC method [32], relative to the pool of RNA containing equal amounts of each study sample. Statistical analysis was performed with the use of statistical package SPSS 17.0.

Gene expression in both study groups was compared by Wilcoxon test (paired comparisons between tumor and normal tissue) and Mann-Whitney/Kruskal-Wallis tests (testing differences between independent groups). Chi-square test was used for comparison of proportions. Microarray data analysis was performed by routinely applied Student’s t-test. In the case of ALK, a threshold for “overexpression” was adopted. The most common activating translocation for ALK in NSCLC – EMLA-ALK – is detected in 1.6–9% of surgical series and amplification of ALK was found in 10% of cases [19]. Therefore, we adopted the 90th percentile expression level for the entire cohort of analyzed NSCLC samples (n=85) as the overexpression threshold, so that samples with the expression from the highest 10 percentiles (n=9) were considered to show “overexpression”. For a gene expressed only in a fraction of samples, i.e. AKR1B10, the positivity threshold was defined as expression above the expression in calibrator sample (pool of RNA).

Metastasis-free survival was calculated with Kaplan-Meier method from pulmonary surgery to the diagnosis of any metastasis, or censored due to end of follow up or death. Survival curves were compared using log-rank test. Owing to the multiple comparisons, a conservative significance threshold of p=0.005 was applied for all statistical analyses except for survival comparisons.

### RESULTS

#### Tumor gene expression in smokers vs. never-smokers

Tumor expression of 5 genes was significantly higher in never-smokers compared to smokers: CSF1R (p=0.0001), RRAD (p<0.0001), PR (p=0.0004), TGFBR2 (p=0.0027) and EPHB6 (p=0.0033), whereas the expression of AR was at the borderline level (p=0.008) (Tab. 3; Fig. 1). The expression of RRAD (p=0.0003), CSF1R (p=0.0008) and PR (p=0.0009) in adenocarcinomas of never-smoking females was higher compared to their smoking counterparts. The expression of CHRNA6 tended to be higher (p=0.023) and the expression of survivin lower (p=0.007) in adenocarcinomas of never-smoking compared to smoking women.

Among 9 patients with overexpression of ALK transcript, 4 were never-smokers, 4 were former-smokers and 1 was an active smoker. This group included 7 females and 2 males;
Table 2. Brief characteristics of the analyzed genes.

<table>
<thead>
<tr>
<th>#</th>
<th>Gene symbol</th>
<th>Assay-on-Demand ID used for analysis</th>
<th>Full gene name</th>
<th>Gene product function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AHR</td>
<td>AHR-Hs00169233_m1</td>
<td>aryl hydrocarbon receptor</td>
<td>cytosolic transcription factor</td>
<td>[12, 40]</td>
</tr>
<tr>
<td>2</td>
<td>AKR1B10</td>
<td>AKR1B10-Hs0025252_m1</td>
<td>aldo-keto reductase family 1, member B10 (aldo reductase)</td>
<td>enzyme reducing aliphatic and aromatic aldehydes</td>
<td>[13, 41, 42]</td>
</tr>
<tr>
<td>3</td>
<td>ALK</td>
<td>ALK-Hs00608289_m1</td>
<td>anaplastic lymphoma kinase</td>
<td>receptor tyrosine kinase; can form fusion genes</td>
<td>[5, 18-20, 48]</td>
</tr>
<tr>
<td>4</td>
<td>AR</td>
<td>AR-Hs00171172_m1</td>
<td>androgen receptor</td>
<td>a DNA-binding transcription factor regulating gene expression</td>
<td>[10, 33, 35]</td>
</tr>
<tr>
<td>5</td>
<td>BIRC5</td>
<td>BIRC5-Hs00978503_m1</td>
<td>baculoviral inhibitor of apoptosis repeat-containing 5, survivin</td>
<td>inhibit caspase activation, thereby leading to negative regulation of apoptosis or programmed cell death</td>
<td>[24, 46]</td>
</tr>
<tr>
<td>6</td>
<td>CDKN2A</td>
<td>CDKN2A-Hs00233365_m1</td>
<td>cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)</td>
<td>tumor suppressor protein, cell cycle regulation</td>
<td>[16, 17]</td>
</tr>
<tr>
<td>7</td>
<td>CHRNA6</td>
<td>CHRNA6-Hs00610231_m1</td>
<td>Cholinergic receptor, nicotinic, alpha 6</td>
<td>regulates cellular proliferation and apoptosis via Akt pathway</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>8</td>
<td>CHRNB3</td>
<td>CHRNB3-Hs00181269_m1</td>
<td>Neuronal acetylcholine receptor subunit beta-3</td>
<td>regulates cellular proliferation and apoptosis via Akt pathway</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>9</td>
<td>CSF1R</td>
<td>CSF1R-Hs00911250_m1</td>
<td>Colony stimulating factor 1 receptor; cell-surface protein</td>
<td>cytokine which controls the production, differentiation, and function of macrophages</td>
<td>[25, 26, 44]</td>
</tr>
<tr>
<td>10</td>
<td>DLG1</td>
<td>DLG1-Hs00938204_m1</td>
<td>Discs, large homolog 1 (Drosophila)</td>
<td>involved in the regulation of cell polarity and proliferation control</td>
<td>[27]</td>
</tr>
<tr>
<td>11</td>
<td>EPHB6</td>
<td>EPHB6-Hs00270052_m1</td>
<td>ephrin type-B receptor 6</td>
<td>receptor for transmembrane protein</td>
<td>[26, 43]</td>
</tr>
<tr>
<td>12</td>
<td>ER1</td>
<td>ER1-Hs00174860_m1</td>
<td>estrogen receptor 1</td>
<td>member of the nuclear hormone family of intracellular receptors; DNA-binding transcription factor that regulates gene expression</td>
<td>[11, 33, 36, 37]</td>
</tr>
<tr>
<td>13</td>
<td>ER2</td>
<td>ER2-Hs00230957_m1</td>
<td>estrogen receptor 2</td>
<td>member of the nuclear hormone family of intracellular receptors; DNA-binding transcription factor that regulates gene expression</td>
<td>[11, 33, 36, 37]</td>
</tr>
<tr>
<td>14</td>
<td>IGF1</td>
<td>IGF1-Hs01547657_m1</td>
<td>Insulin-like growth factor 1 tyrosine kinase</td>
<td>polypeptide protein hormone similar in molecular structure to insulin</td>
<td>[21]</td>
</tr>
<tr>
<td>15</td>
<td>IGF1R</td>
<td>IGF1R-Hs00609566_m1</td>
<td>insulin-like growth factor 1 receptor</td>
<td>transmembrane receptor that is activated by IGF-1 and by the related growth factor IGF-2</td>
<td>[22]</td>
</tr>
<tr>
<td>16</td>
<td>IGF2</td>
<td>IGF2-Hs01005963_m1</td>
<td>Insulin-like growth factor 2 tyrosine kinase</td>
<td>protein hormone sharing structural similarity to insulin</td>
<td>[23]</td>
</tr>
<tr>
<td>17</td>
<td>PR</td>
<td>PR-Hs00172183_m1</td>
<td>progesterone receptor</td>
<td>intracellular steroid receptor specifically binding progesterone</td>
<td>[9, 33, 34]</td>
</tr>
<tr>
<td>18</td>
<td>RAD</td>
<td>RAD-Hs00188163_m1</td>
<td>Ras-related associated with diabetes</td>
<td>GTP-binding protein RAD</td>
<td>[28, 29]</td>
</tr>
<tr>
<td>19</td>
<td>SOX9</td>
<td>SOX9-Hs00165814_m1</td>
<td>SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal)</td>
<td>important role in respiratory epithelium development; has an impact on AR protein expression</td>
<td>[30, 47]</td>
</tr>
<tr>
<td>20</td>
<td>TGFBR2</td>
<td>TGFBR2-Hs00234253_m1</td>
<td>transforming growth factor, beta receptor II</td>
<td>polypeptide growth factor</td>
<td>[7]</td>
</tr>
<tr>
<td>21</td>
<td>TGFBR3</td>
<td>TGFBR3-Hs00234257_m1</td>
<td>transforming growth factor, beta receptor III</td>
<td>polypeptide growth factor</td>
<td>[7]</td>
</tr>
</tbody>
</table>
Table 3. Differences in expression of analyzed genes between smokers and non-smokers. Significant differences are marked in bold letters.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression in tumors of never-smokers and smokers</th>
<th>Expression in tumors and normal lungs of never-smokers</th>
<th>Expression in tumors and normal lungs of smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fold change (median expression)</td>
<td>Fold change (median expression)</td>
<td>Fold change (median expression)</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>p value</td>
<td>p value</td>
</tr>
<tr>
<td>AHR</td>
<td>1.16 (0.89/0.77); 0.0077</td>
<td>1.09 (0.89/0.82); 0.5</td>
<td>0.86 (0.773/0.897); &lt;0.0001</td>
</tr>
<tr>
<td>ARK1B10</td>
<td>0.23 (1.43/6.14); 0.9091</td>
<td>23.8 (1.43/0.06); &lt;0.0001</td>
<td>458.32 (6.1415/0.034); &lt;0.0001</td>
</tr>
<tr>
<td>ALK</td>
<td>1.12 (2.61/2.34); 0.3541</td>
<td>3.39 (2.61/0.77); 0.105</td>
<td>3.82 (2.33/0.61); 0.601</td>
</tr>
<tr>
<td>AR</td>
<td>1.57 (0.74/0.47); 0.0078</td>
<td>0.53 (0.74/1.39); &lt;0.0001</td>
<td>0.38 (0.47/1.25); &lt;0.0001</td>
</tr>
<tr>
<td>BIRC5</td>
<td>0.45 (0.91/2.03); 0.0091</td>
<td>5.35 (0.91/1.17); &lt;0.0001</td>
<td>12.0 (2.03/0.17); &lt;0.0001</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>1.17 (2.60/2.22); 0.9454</td>
<td>18.57 (2.60/0.14); 0.011</td>
<td>11.94 (2.22/0.19); &lt;0.0001</td>
</tr>
<tr>
<td>CHRNA6</td>
<td>1.22 (2.81/2.30); 0.0346</td>
<td>6.39 (2.81/0.44); &lt;0.0001</td>
<td>3.24 (2.30/0.71); &lt;0.0001</td>
</tr>
<tr>
<td>CHRNA8</td>
<td>2.35 (2.19/0.93); 0.0990</td>
<td>3.59 (2.19/0.61); 0.36</td>
<td>0.45 (0.93/0.06); 0.816</td>
</tr>
<tr>
<td>CSF1R</td>
<td>1.89 (0.72/0.38); &lt;0.0001</td>
<td>1.29 (0.72/0.56); 0.181</td>
<td>0.63 (0.38/0.60); &lt;0.0001</td>
</tr>
<tr>
<td>Dlg1</td>
<td>1.04 (0.76/0.73); 0.5080</td>
<td>0.86 (0.76/0.88); 0.028</td>
<td>0.85 (0.73/0.86); 0.014</td>
</tr>
<tr>
<td>EPHB6</td>
<td>2.06 (0.99/0.48); 0.0033</td>
<td>0.63 (0.99/1.58); &lt;0.0001</td>
<td>0.31 (0.48/1.55); &lt;0.0001</td>
</tr>
<tr>
<td>ERI</td>
<td>0.95 (1.30/1.37); 0.7459</td>
<td>1.38 (1.30/0.94); 0.039</td>
<td>1.47 (1.37/0.93); 0.007</td>
</tr>
<tr>
<td>ER2</td>
<td>1.07 (1.62/1.52); 0.8230</td>
<td>1.69 (1.62/0.96); 0.029</td>
<td>1.6 (1.52/0.95); 0.027</td>
</tr>
<tr>
<td>IGF1</td>
<td>1.42 (1.38/0.97); 0.9600</td>
<td>1.77 (1.38/0.78); 0.914</td>
<td>1.11 (0.970/8.7); 0.97</td>
</tr>
<tr>
<td>IGF1R</td>
<td>0.80 (0.61/0.76); 0.1153</td>
<td>0.82 (0.61/0.74); 0.031</td>
<td>1.04 (0.76/0.73); 0.509</td>
</tr>
<tr>
<td>IGF2</td>
<td>0.79 (0.27/0.34); 0.7321</td>
<td>1.5 (0.27/0.18); 0.003</td>
<td>2.19 (0.339/0.155); 0.062</td>
</tr>
<tr>
<td>PR</td>
<td>2.11 (0.40/0.19); 0.0004</td>
<td>0.39 (0.40/1.02); &lt;0.0001</td>
<td>0.18 (0.19/1.03); &lt;0.0001</td>
</tr>
<tr>
<td>RRAD</td>
<td>2.75 (0.55/0.20); &lt;0.0001</td>
<td>0.65 (0.55/0.85); 0.025</td>
<td>0.15 (0.20/1.30); &lt;0.0001</td>
</tr>
<tr>
<td>SOX9</td>
<td>1.08 (1.28/1.18); 0.3933</td>
<td>3.05 (1.28/0.42); &lt;0.0001</td>
<td>2.62 (1.18/0.45); 0.002</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>1.52 (0.41/0.27); 0.0027</td>
<td>0.42 (0.41/0.97); &lt;0.0001</td>
<td>0.28 (0.27/0.97); &lt;0.0001</td>
</tr>
<tr>
<td>TGFBR3</td>
<td>1.83 (0.22/0.12); 0.0275</td>
<td>0.19 (0.22/1.17); &lt;0.0001</td>
<td>0.13 (0.12/0.92); &lt;0.0001</td>
</tr>
</tbody>
</table>

Figures to Table 3 available from authors by request.

8 adenocarcinomas and 1 squamous cell carcinoma. Four of these patients developed distant metastases including 2 brain metastases. Overexpression of ALK was also found in 7 normal lung specimens. Interestingly, ALK overexpression coincided in lung and tumor tissue in only one case. If 75th percentile of expression was applied as a threshold for overexpression, ALK was found in 8 normal lung samples, with no expression in the corresponding tumor sample; 7 cases exhibited overexpression in both tumor and the corresponding normal lung and in 14 cases ALK was overexpressed in NSCLC and not in the corresponding normal lung.

AKR1B10 was expressed in 23 NSCLCs, in a similar proportion of smokers (31%) and never-smokers (19%; p=0.31).

Gene expression in tumor samples and in corresponding normal lung tissue in smokers vs. never-smokers

Expression of 7 genes was significantly lower in tumors than in normal lung tissue: AR (p<0.0001), EPHB6 (p<0.0001), PR (p<0.0001), TGFBR2 (p<0.0001), TGFBR3 (p<0.0001), ER1 (p=0.0006) and Dlg1 (p=0.0016). When the gene expression in tumor samples and corresponding normal lung tissue was analyzed separately for smokers and never-smokers, there was no statistically significant difference in gene expression for ERI1 (p=0.007 and p=0.039, respectively) and Dlg1 (p=0.014 and p=0.028, respectively). Tumor expression of three genes: AHR (p<0.0001), CSF1R (p<0.0001) and RRAD (p<0.0001) was lower than in healthy lung tissue in smokers but not in never-smokers.

The expression of 6 genes was significantly higher in tumors compared to normal lung tissue: AKR1B10 (p<0.0001), CDKN2A (p<0.0001), CHRNA6 (6-fold in never-smokers and 3-fold in smokers; p<0.0001), SOX9 (p<0.0001), survivin (p<0.0001) and ER2 (p=0.002). Mean expression of IGF2 was higher in tumor compared to healthy lung tissue in never-smokers (p=0.003) but not in smokers. The fold change in the expression level of ER2 in tumor samples and in corresponding normal lung tissue both in smokers and never-smokers was above the established threshold of p=0.005 (p=0.027 and p=0.029, respectively).
Irrespective of smoking status, no expression of AKR1B10 was found in normal lung tissue. All other genes were expressed at the same levels in normal lungs of smokers and never-smokers. In the analysis confined to women, expression of CDKN2A was higher (p=0.0013) in normal lung of smoking patients. Intriguingly, although not significant, an increased expression of lung RRAD was found in 20% of smoking women. The expression of SOX9 was slightly higher (p=0.06) in lungs of smoking women, whereas the expression of TGFBR3 tended to be lower in normal lung of smoking females (p=0.0051).

Prognostic impact of gene expression

There was no difference in metastasis-free survival between smokers and never-smokers (p=0.74). ER2 (p=0.01) and AKR1B10 (p=0.013) were significantly correlated with metastasis-free survival in the entire cohort. However, none of the analyzed genes was correlated with metastasis-free survival after correction for multiple comparisons.

DISCUSSION

This study showed distinct tumor expression of several genes potentially involved in NSCLC carcinogenesis in never-smoking compared to smoking patients, whereas none of the analyzed genes showed different expression in normal lung of smokers compared to never-smokers. This suggests that the differences in tumor expression between smokers and never-smokers reflect activation of smoking-related pathways in NSCLC patients rather than a mere impact of cigarette smoke insult on gene expression in pulmonary epithelium. The same holds true in the analysis restricted to women, which provides a strong suggestion that expression of sex hormone receptors in the normal lung is not affected by smoking status.

Our data suggest that aberrant sex hormone receptor signaling may be among the causes of NSCLC in both smokers and never-smokers. As in the study of Marquez-Garban [33], expression of PR and AR was lower in NSCLC than in normal lung. As progesterone was shown to inhibit the growth of NSCLC cell lines with expression of PR in vitro and in vivo [34], low expression of PR in tumors may be viewed as a means of averting the antiproliferative effect of progesterone by NSCLC. Similarly, androgens (e.g. fluoxymesterone) were demonstrated to inhibit the growth of breast cancer through the activation of AR and consequent induction of apoptosis [35]. It may be speculated therefore that the administration of PR or AR agonists might provide clinical benefit in both smoking and non-smoking NSCLC patients [34].

Importantly, the tumor expression of PR was significantly lower in smokers than in never-smokers both in all patients and in women. A similar trend was noted for AR. This finding might suggest that biology of NSCLC in smokers involves a higher degree of intolerance to inhibitory effects of progesterone. It was previously shown that carcinogens from cigarette smoke (BaP, benzo[a]pyrene) may disrupt protective...
AHR is a transcriptional factor involved in detoxification responses to pollutants and in intrinsic biological processes of multicellular organisms. AHR, upon activation by benzopyrenes (BPs), translocates into the nucleus and induces expression of several genes, including CYP1A1, that further activate BPs’ carcinogenic potential. It was also previously shown that AHR can shuttle polycyclic aromatic hydrocarbons to the nucleus and enhance oxidative DNA damage [44]. Relatively lower expression of AHR in tumors of smokers may suggest their intolerance to undue genetic insult. AKR1B10 was expressed in a proportion of tumors of both smokers and never-smokers. Reduction of benzopyrenes by AKR1B10 in lung cancer cells leads to ROS-mediated genotoxicity and contributes to lung carcinogenesis. The lack of expression of this enzyme in normal lung of smokers suggests that upregulation of AKR1B10 may represent a specific tumor-promoting aberration [41]. Experimental data suggest that inhibitors of AKR1B10 can enhance the chemopreventive effect of retinoids [42].

In our series, expression of EPHB6, a kinase receptor, was significantly lower in tumors compared to healthy lung tissue, and its tumor expression was significantly higher in non-smokers compared to smokers. This gene is frequently silenced by promoter hypermethylation [43] and its low expression was found to correlate with increased risk of metastasis in breast cancer. This may suggest a potential therapeutic role of demethylating agents in these patients.

We confirmed the previously reported downregulation of transforming growth factor receptors, TGFBR2 and TGFBR3, in transition from normal lung to NSCLC [7]. However, in our series low expression of these genes was also found in NSCLCs of smokers, possibly through hypermethylation, as an early event in carcinogenesis. Further, tumor expression of TGFBR2 was significantly lower in smokers compared to never-smokers. It may be envisaged that NSCLC carcinogenesis in both smokers and never-smokers is dependent on silencing of antigrowth effect of TGF, albeit to a different extent. The downregulation of EPHB6, TGFBR2 and TGFBR3 was confirmed in silico validation of our results. In one of the datasets TGFBR3 gene was confirmed to be a very sensitive malignancy marker (a fold change of 0.18).

Tumor expression of CSF1R, a colony stimulating factor for macrophage receptor, in smokers was significantly lower compared to normal lung and to tumors of never-smokers. In cell cultures, CSF1R was shown to immortalize the epithelial cells through upregulation of telomerase [44]. Upregulation of CSF1 is correlated with poor prognosis in squamous cell and in other types of lung cancer [25]. In our study, tumor CSF1R expression was dependent on the smoking status. This finding requires further investigation, as the biological role of macrophages (growth promoting and inhibiting) is tumor dependent [45]. Agents affecting CSF1R activity could have potential therapeutic effect.
Tumor expression of survivin and SOX9, a transcription factor, was higher in tumors compared to healthy lung tissue, and this difference was more pronounced in smokers. The expression of SOX9 was insignificantly higher in lungs of smokers compared to lungs of never-smokers, suggesting its role in early carcinogenesis in smokers. Survivin inhibitors are currently being tested in clinical trials [46]. A relatively high expression of SOX9 in adenocarcinomas of smoking women as compared to their never-smoking counterparts might guide patient selection in future studies. In colorectal cancer, the down-regulation of SOX9 inhibited cell proliferation in vitro and the growth of tumor xenografts [47]. Therefore, SOX9 gene might also be a potential target for future therapies.

IGF2 expression was increased in tumors compared to healthy lung tissue in never-smoking NSCLC patients. This gene is commonly overexpressed in many human cancers, e.g. breast, colon or prostate cancer, and its higher expression has been attributed to a loss of imprinting [23]. Biological significance of these findings warrants further studies.

ALK overexpression was detected in NSCLCs of both smokers and never-smokers and was not tumor-specific. This is an interesting finding, suggesting that ALK overexpression may be among early events in lung carcinogenesis, however a question of whether the transcript found in normal lung of cancer patients is canonical or one of the fusion variants, remains to be established. Similarly to our findings, Martelli et al. [48] found expression of fusion gene EML4-ALK in 15% of histologically normal lung samples, with no corresponding tumoral expression of this fusion transcript in the same patients.

To validate the obtained results, the expression of relevant genes was examined in the publicly available gene expression microarray datasets. First, we evaluated genes expressed in pulmonary epithelium as a response to tobacco smoke insult. The potentially relevant dataset GSE23546 [49] included 1349 samples of non-tumorous human lung tissue, hybridized on Rosetta/Merck Human RSTA Custom Affymetrix 2.0 microarrays. However, GEO data repository did not include clear annotation of the smoking status for the individual cases. The data set GSE7895 [50] describes gene expression of 104 NSCLC patients: 21 never-smokers, 52 current smokers and 31 former smokers. This experiment was carried on Affymetrix HG U133a microarrays and covered expression of CSF1R, RRAD, TGFB2, EPHB6, AR, and CHRNA6. Three genes (RRAD, TGFB2 and EPHB6) were confirmed to be downregulated in current smokers (Appendix A: Fig. 1S-3S; Appendix B: Tab. 1S). Importantly, for RRAD gene both probes indicated the difference in expression with high significance (Student’s t-test; p<0.0005). We also investigated the differences in gene expression in cancer vs. healthy lung tissue in two datasets published in GEO. The dataset GSE12428 [51] includes information on gene expression of 62 patients: 31 lung cancer and 31 non-tumorous tissues. Experiments were performed with Agilent two channel (Cy3/Cy5) Whole Human Genome Oligo Microarrays, design number 012391 (GPL1708). In this dataset we were able to evaluate 7 genes: AR, EPHB6, PR (PGR), TGFB2 (TGFB1), TGFB3, ER1 (ESR1) and DLG1, three of which were confirmed to be downregulated in tumors: PR (p=1.78E-21), TGFB2 (p=0.009) and TGFB3 (p=4.70E-11) (Appendix A: Fig. 4S; Appendix B: Tab. 2S). Other genes did not show significant expression differences. Another database GSE10072 [52] includes gene expression information obtained from 107 samples: 58 lung adenocarcinomas and 49 normal lung tissues. Tissue mRNAs were hybridized to Affymetrix HG U133a microarrays, therefore for many genes of interest several probe sets were tested. Data confirmed TGFB3 gene to be the most significantly downregulated in tumors vs. normal lung with a fold change of 0.18 (p=0.18E-34). The results for all other genes are included in the supplemental files, together with appropriate figures (Appendix A: Fig. 5S-8S; Appendix B: Tab. 3S).

CONCLUSIONS

In conclusion, we found significant differences in tumor expression of several genes between smoking and never-smoking NSCLC patients. Some of these differences (e.g. related PR, CSF1R) have not been previously reported, whereas others (for RRAD, TGFB2, EPHB6) were confirmed. Downregulation of PR and upregulation of ER2 in tumors compared to normal lungs suggests possible hormonal dependence of NSCLC in both smokers and never-smokers. Thus, therapeutic potential of progesterone and estrogen inhibitors warrants further investigations. AKR1B10 is not expressed in lung parenchyma, which makes it a potentially highly specific, however moderately sensitive, NSCLC marker. ALK overexpression seems not to be lung cancer specific. Distinct molecular features of NSCLC of never-smokers (e.g. CHRNA6 upregulation) may prompt new treatment strategies.

ACKNOWLEDGEMENTS

This study was financially supported by National Science Centre, Poland, grant no. N402 238137.

Conflict of interest

The authors declare no conflict of interest.
REFERENCES


