An integrative strategy for targeted evaluation of biomarker expression in non-small cell lung cancer

JOHANNA MATTSSON
Abstract

Despite improvements in therapy, the prognosis for non-small cell lung cancer (NSCLC) patients remains poor, and cure is only possible in localized tumors after surgical resection. A new generation of targeted cancer drugs has led to the expectation that lung cancer therapy can be significantly improved, but these drugs are today only an option in a small subset of NSCLC patients, and their effect is temporary. Therefore, the aim of this thesis was to characterize NSCLC in order to find new treatment targets and to evaluate biomarkers that further optimize therapy selection.

In Paper I, the expression of the potential treatment targets claudin 6 and claudin 18.2 were evaluated based on immunohistochemical- and gene expression analysis. High ectopic protein and gene expression were demonstrated for both claudins in small subgroups of NSCLC. Clinical trials using humanized monoclonal antibodies against both proteins are ongoing in other cancer forms and may be extended to NSCLC.

In Paper II, the prognostic impact of the inflammatory mediator cyclooxygenase 2 (COX-2) was evaluated. No prognostic significance was found in a meta-analysis incorporating gene expression data of 1337 NSCLC patients. Likewise, COX-2 protein expression in tumor cells was not associated with survival in two independent NSCLC cohorts. However, in one of the analyzed cohorts, higher COX-2 expression in the tumor stroma was associated with longer survival and may therefore be a subject for further investigation.

In Paper III, tumor and stromal COX-2 protein expression was examined in patients treated with the COX-2 inhibitor celecoxib in order to evaluate if COX-2 expression is a predictive biomarker for benefit of celecoxib therapy. Celecoxib did not prolong overall survival neither in the whole cohort nor in patients stratified according to COX-2 expression in tumor or stromal cells. Noteworthy, a tendency towards longer survival was again demonstrated in patients with high COX-2 stromal expression.

In Paper IV, the diagnostic methods for identification of ALK rearrangements were assessed in a large representative Swedish NSCLC population. Fluorescence in situ hybridization (FISH), as the diagnostic standard, was compared to two immunohistochemical assays. ALK gene expression levels were incorporated to supplement the molecular data. The frequency of ALK rearrangements was lower than previously reported. The different methods to detect the ALK fusion demonstrated overlapping results. However, the overlap was poor, so the methods cannot be regarded as interchangeable and should thereby be interpreted with caution when used in clinical diagnostics.

In summary, this thesis applied an integrative translational approach to characterize potential new treatment targets and to evaluate the detection of existing predictive biomarkers in NSCLC.

Keywords: non-small cell lung cancer, prognostic biomarkers, predictive biomarkers, immunohistochemistry, gene expression, COX-2, claudin, ALK

Johanna Mattsson, Department of Immunology, Genetics and Pathology, Molecular and Morphological Pathology, Rudbecklaboratoriet, Uppsala University, SE-751 85 Uppsala, Sweden.

© Johanna Mattsson 2016

ISSN 1651-6206
urn:nbn:se:uu:diva-285844 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-285844)
Till mamma och pappa
Supervisors and Examining Board

Main supervisor  Patrick Micke, M.D., Associate Professor
Department of Immunology, Genetics and Pathology
Uppsala University, Uppsala

Assistant supervisors  Johan Botling, M.D., Associate Professor
Department of Immunology, Genetics and Pathology
Uppsala University, Uppsala

Fredrik Pontén, M.D., Professor
Department of Immunology, Genetics and Pathology
Uppsala University, Uppsala

Faculty opponent  Åslaug Helland, M.D., Associate Professor
Department of Cancer Genetics
Oslo University, Oslo

Examinining committee  Marie-Louise Fjällskog, M.D., Associate Professor
Department of Medical Sciences, Endocrine Oncology
Uppsala University, Uppsala

Marika Nestor, M.D., Associate Professor
Department of Immunology, Genetics and Pathology
Uppsala University, Uppsala

Katalin Dobra, M.D., Associate Professor
Department of Laboratory Medicine
Karolinska Institute, Stockholm
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


*Contributed equally to the work.

Reprints were made with permission from the respective publishers.
Related publications


## Contents

Introduction ................................................................................................... 13
Lung cancer .............................................................................................. 14
  Epidemiology.......................................................................................... 14
  Etiology ................................................................................................. 14
  Histology and differences at the molecular level ..................................... 15
  Diagnosis and staging ......................................................................... 18
Treatment ............................................................................................... 20
  Prognostic and predictive biomarkers .................................................. 23
Immunohistochemistry (IHC) ................................................................. 24
Fluorescence in situ hybridization (FISH) .............................................. 26
Tissue microarray (TMA) ......................................................................... 29
Gene expression microarray ..................................................................... 30

Present Investigation ..................................................................................... 32
  Aim ......................................................................................................... 32
  Molecular targets studied ....................................................................... 32
    Claudin 6 and claudin 18.2 ................................................................. 32
    Cyclooxygenase 2 (COX-2) ................................................................. 33
    Anaplastic lymphoma kinase (ALK) .................................................. 33
Material and Methods ............................................................................... 34
  Patient cohorts and study design ......................................................... 34
  Methodological considerations ............................................................ 38
Results and discussion .............................................................................. 40
  Paper I: ................................................................................................. 40
  Paper II: ............................................................................................... 41
  Paper III: .............................................................................................. 42
  Paper IV: .............................................................................................. 44

Concluding remarks and future perspectives ............................................. 46

Acknowledgements ....................................................................................... 48
References ..................................................................................................... 50
Abbreviations

ABC  Avidin-biotin complex
ALK  Anaplastic lymphoma kinase
AP   Alkaline phosphatase
BRAF v-raf murine sarcoma viral oncogene homolog B
cDNA Complementary DNA
cRNA Complementary RNA
cTNM Clinical tumor, node, metastasis
CK5  Cytokeratin 5
COX  Cyclooxygenase
CT   Computed tomography
DAB  3,3'-diaminobenzidine
DAPI 4',6-diamidino-2-phenylindole
DDR2 Discoidin domain-containing receptor 2
DNA  Deoxyribonucleic acid
EGFR Epidermal growth factor receptor
EML4 Echinoderm microtubule associated protein like 4
ERBB2 v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2
FDA  Food and Drug Administration
FFPE Formalin-fixed paraffin embedded
FGFR1 Fibroblast growth factor receptor 1
FISH Fluorescence in situ hybridization
H2O2 Hydrogen peroxide
HE  Hematoxylin and eosin
HQ  3-hydroxy-2-quinoxaline
HR  Hazard ratio
HRP Horseradish peroxidase
IASLC International Association for the Study of Lung Cancer
Ig  Immunoglobulin
IHC Immunohistochemistry
ISH In situ hybridization
KEAP1 Kelch-Like ECH-Associated Protein 1
KRAS Kirsten rat sarcoma viral oncogene homolog
mRNA messenger RNA
n/a Not applicable
NOS Not otherwise specified
NSAID Non-steroidal anti-inflammatory drugs
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSCLC</td>
<td>Non-small cell lung cancer</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PD1</td>
<td>Programmed cell death protein 1</td>
</tr>
<tr>
<td>PD-L1</td>
<td>Programmed cell death ligand 1</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatase and tensin homolog</td>
</tr>
<tr>
<td>PTGS2</td>
<td>Prostaglandin-endoperoxide synthase 2</td>
</tr>
<tr>
<td>pTNM</td>
<td>Pathological tumor, node, metastasis</td>
</tr>
<tr>
<td>RET</td>
<td>Ret proto-oncogene</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>ROS1</td>
<td>ROS proto-oncogene 1, receptor tyrosine kinase</td>
</tr>
<tr>
<td>SCLC</td>
<td>Small-cell lung cancer</td>
</tr>
<tr>
<td>STK11</td>
<td>Serine/threonine kinase 11</td>
</tr>
<tr>
<td>TKI</td>
<td>Tyrosine kinase inhibitor</td>
</tr>
<tr>
<td>TMA</td>
<td>Tissue microarray</td>
</tr>
<tr>
<td>TNM</td>
<td>Tumor, node, metastasis</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumor protein p53</td>
</tr>
<tr>
<td>TTF1</td>
<td>Thyroid transcription factor 1</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Introduction

Lung cancer is the leading cause of cancer death in Sweden, with more deaths each year than breast and colon cancer combined [1-2]. The overall prognosis of lung cancer is poor [3] and treatment options are limited. Cure is only possible in early stages when surgical resection is an option [2]. Chemotherapy remains the standard treatment for the majority of patients, but response is limited and the impact on survival is modest [4]. During the last decade, treatment options have progressed beyond chemotherapy and the discovery of targetable oncogenic driver mutations has revolutionized the treatment for a small subgroup of advanced lung cancer patients [5-6]. More recently, immunotherapy has emerged as a new potential treatment strategy, demonstrating long term responses in patients with late stage disease [7]. Unfortunately, these new treatment alternatives only show benefit in a subset of patients. Thus, the identification of other therapy targets or biomarkers that may refine the selection of existing treatment modalities are still warranted.

This thesis is based on four papers. An integrative translational approach has been adopted to describe and identify new potential treatment targets as well as to evaluate predictive biomarkers in non-small cell lung cancer (NSCLC).
Lung cancer

Epidemiology

Lung cancer is the most common cancer type worldwide, both regarding the incidence as well as the number of deaths. In 2012, 1.8 million (12.9% of total) were diagnosed and 1.6 million (19.4% of total) died of this disease. Lung cancer was the most common cancer type among men, with 1.2 million new cases (16.7%) and almost 1.1 million deaths (23.6%), whereas in women the incidence was lower, with 583 000 new cases (8.7%) and 491 000 deaths (13.8%) [8]. Lung cancer rates vary by ethnicity, across countries, and between males and females within each country, mainly due to smoking trends [9]. In 2014 in Sweden, lung cancer was the fifth most common cancer type in men and the fourth most common cancer type in women, but as in the rest of the world, it was the most common cause of cancer-related death. In 2014, 4063 patients were diagnosed with lung cancer and 3647 patients died from the disease [2]. The prognosis for patients diagnosed with lung cancer remains poor, with a 5-year survival rate of 13.6% among men and 19.4% among women across all stages [3].

Etiology

Cigarette smoking is the main cause of lung cancer [10], and in Sweden, approximately 90% of all lung cancer cases are due to smoking [2]. Smokers have a 20-fold higher risk of developing lung cancer compared to never smokers and the risk increases with the duration of smoking as well as with the number of cigarettes smoked per day. By ceasing smoking, the risk of developing lung cancer decreases, but former-smokers still have a higher risk of developing lung cancer than never-smokers [11]. Although smoking is the most common cause of lung cancer, never-smokers still suffer from this type of cancer. Globally, about 300,000 lung cancer deaths per year are non-smoking related. The lung cancer risk in never-smokers varies geographically, with the highest risk observed in Asia [12]. Lung cancer in non-smokers is most often of the adenocarcinoma histology, more common in women [13] and in younger patients (<60 years) [14]. In addition to smoking, other known risk factors include exposure to ionizing radiation e.g. radon gas, as well as exposure to occupational agents such as asbestos, chromium, nickel and arsenic. Furthermore, indoor and outdoor air pollution, pre-existing chronic lung disease, lack of physical activity, dietary factors, and second-hand smoke exposure are also associated with an increased lung cancer risk [11]. However, far from everyone who is exposed to cigarette smoke, or to other environmental risk factors, develop lung cancer. The risk of developing lung cancer most likely depends on genetic predisposition together with exposure to respiratory carcinogenic factors [15].
Histology and differences at the molecular level

Lung cancer is a heterogeneous disease, both at the molecular and the histological level. Historically, lung cancer is divided into two main histological subtypes: small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). These two subtypes differ in their biological behavior as well as in treatment and prognosis [16].

SCLC accounts for 10-15% of all lung cancer cases and is believed to originate from neuroendocrine cells in the lung [17]. This type of lung cancer is predominantly caused by substantial cigarette consumption and, compared to NSCLC, is characterized by a faster growth pattern, earlier spread, and higher sensitivity to chemotherapeutic agents. Although SCLC initially responds well to chemotherapy, resistance to treatment is developed and SCLC remains fatal for the majority (95%) of patients [18-19].

NSCLC, which is the focus of this thesis, is the most common type of lung cancer, representing approximately 85% of all lung cancer cases [20]. Based on tumor histology, NSCLC is further sub-classified into three main subgroups: adenocarcinoma, squamous cell carcinoma and large cell carcinoma [21]. Previously, there was no need to differentiate between these histological subtypes when considering the choice of treatment. However, this changed after the recognition that patients with non-squamous histology responded better to certain cytotoxic drugs as well as harbored distinct mutations applicable for targeted therapy [22].

The histological classification is based on an expert recommendation of pathologists appointed by the World Health Organization (WHO). The studies in this thesis applied the histological classification based on the WHO classification of lung tumors from 2004 [21].

Adenocarcinoma

Adenocarcinoma is the largest histological subgroup, comprising approximately 40% of all lung cancer cases in Sweden [20]. Adenocarcinomas are epithelial tumors characterized by a glandular structure, pneumocyte marker expression (e.g. thyroid transcription factor 1 (TTF1) and napsin A) or mucin production. They most often occur in the lung periphery [23], where they are derived from small bronchi, bronchioles and alveoli [24]. Recently the WHO has proposed a further subclassification of invasive adenocarcinoma with regard to the architectural structures of acinar, papillary, micropapillary, lepidic or solid patterns [23]. These subgroups are of relevance regarding prognosis, but their potential role in guiding treatment decisions and surgical management is still not clear [25]. Among never-smokers, adenocarcinoma is the most frequent histological type [12]. Moreover, a
higher frequency of adenocarcinomas has been observed in patients from Southern and Eastern Asia compared to other parts of the world [24].

The histological subtypes also differ at the molecular level. In adenocarcinoma patients, mutations are commonly found in Kirsten rat sarcoma viral oncogene homolog (KRAS), epidermal growth factor receptor (EGFR), v-raf murine sarcoma viral oncogene homolog B (BRAF) and v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2). Rearrangements of anaplastic lymphoma kinase (ALK), ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) and ret proto-oncogene (RET) are also common [26]. Furthermore, mutations are frequently observed in the tumor suppressor genes tumor protein P53 (TP53), kelch-like ECH-associated protein 1 (KEAP1) and serine/threonine kinase 11 (STK11) [27]. The frequencies of common genomic aberrations in adenocarcinoma are outlined in Table 1.

Squamous cell carcinoma

Squamous cell carcinoma comprises approximately 20% of all lung cancer cases in Sweden [20]. During the first half of the 20th century, squamous cell carcinoma was the most common subtype of NSCLC, but due to decreased smoking consumption as well as changes in cigarette filters and composition, a decline in patients diagnosed with this subtype of lung cancer has occurred [29,30]. Tumors of squamous cell carcinoma histology are characterized by intracellular bridges and/or keratin structures. The tumors originate from bronchial epithelial cells and are often centrally located in the lung, i.e. in the main or lobar bronchus [21]. The expression of the immunohistochemical markers p40, p63, cytokeratin 5 (CK5) and cytokeratin 5/6 (CK5/6) are commonly used to characterize this subtype of NSCLC [23].

Squamous cell carcinomas display a very high mutation rate, reflecting the carcinogenic effects of tobacco usage [28,30]. Common genetic alterations in squamous cell carcinoma include mutations in the tumor suppressor genes TP53 [31] and phosphatase and tensin homolog (PTEN). Other common aberrations include mutations in discoidin domain-containing receptor 2 (DDR2) and phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA), as well as amplifications of fibroblast growth factor receptor 1 (FGFR1) [26]. Frequencies of common genomic aberrations in squamous cell carcinoma are outlined in Table 1.
Table 1. *Common genomic aberrations in NSCLC.*

<table>
<thead>
<tr>
<th>Adenocarcinoma</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRAS</td>
<td>25-30%</td>
</tr>
<tr>
<td>EGFR</td>
<td>15-20%</td>
</tr>
<tr>
<td>BRAF</td>
<td>1-3%</td>
</tr>
<tr>
<td>TP53</td>
<td>46%</td>
</tr>
<tr>
<td>KEAP1</td>
<td>17%</td>
</tr>
<tr>
<td>STK11</td>
<td>17%</td>
</tr>
<tr>
<td>ERBB2</td>
<td>1-3%</td>
</tr>
<tr>
<td>ALK</td>
<td>3-7%</td>
</tr>
<tr>
<td>ROS1</td>
<td>1-3%</td>
</tr>
<tr>
<td>RET</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Squamous cell carcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP53</td>
</tr>
<tr>
<td>PIK3CA</td>
</tr>
<tr>
<td>PTEN</td>
</tr>
<tr>
<td>DDR2</td>
</tr>
<tr>
<td>FGFR1</td>
</tr>
</tbody>
</table>

**Large-cell carcinoma**

Large-cell carcinomas constitute approximately 10% of all lung cancer cases in Sweden [20]. These tumors lack histological features of squamous cell carcinoma or adenocarcinoma differentiation [21]. However, based on immunohistochemical and mucin staining, the majority of large cell carcinomas can today be allocated to the adenocarcinoma or squamous cell carcinoma subgroup [32]. Only cases that are negative with additional staining, or when staining is unavailable, are confined to the large cell carcinoma subgroup. Therefore, the proportion of diagnosed large cell carcinomas is decreasing [33]. Not surprisingly, mutations in large cell carcinoma are dispersed between adenocarcinoma and squamous cell carcinoma clusters and likewise, based on molecular features, large cell carcinomas can be allocated to the adenocarcinoma or squamous cell cancer group [34].

In addition to these three major sub-categories, other very rare entities of NSCLC exist, e.g. adenosquamous carcinoma and sarcomatoid carcinoma [23].
In the studies included in this thesis, all entities that were not assigned to the adenocarcinoma or squamous cell carcinoma subgroup were considered as large cell carcinoma, not otherwise specified (NOS).

**Diagnosis and staging**

Patients presenting symptoms indicative of lung cancer are initially referred to a chest x-ray. If cancer indication remains, a computed tomography (CT) scan and positron emission tomography (PET) to assess the tumor extension is performed [35]. To ultimately confirm the lung cancer diagnosis, a histological or cytological analysis is necessary [35-36]. The tissue or cell specimen is obtained by either a bronchoscopy or, if the tumor is located peripherally, a transthoracic puncture biopsy. To exclude mediastinal lymph node metastases, a mediastinoscopy, endobronchial ultrasound or endoscopic ultrasound is performed [35].

Staging of lung cancer is important in order to determine the individual cancer burden in a patient and to guide treatment decisions. The staging system applied in most types of cancer is the TNM system, which describes the size of the primary tumor and whether it has invaded nearby tissue (T), the presence of lymph node metastasis (N) and the presence of distant metastasis (M) [37] (Table 2). Currently, the 7th edition of the TNM system is in use [38]. The TNM staging in NSCLC can be based on the clinical or radiological extent of the cancer, referred to as clinical TNM (cTNM), or after surgery based on the pathological findings, referred to as pathological TNM (pTNM) [39].
Table 2. TNM classification of carcinomas of the lung. Reproduced from [38].

<table>
<thead>
<tr>
<th>Primary tumor (T)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>Tumor ≤ 3 cm, surrounded by lung or visceral pleura, no bronchoscopic evidence of invasion more proximal than the lobar bronchus</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumor ≤ 2 cm</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumor &gt; 2 cm but ≤ 3 cm</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor &gt; 3 cm but ≤ 7 cm or tumor with any of the following:</td>
</tr>
<tr>
<td></td>
<td>- Involvement of main bronchus, ≥ 2 cm distal to the carina</td>
</tr>
<tr>
<td></td>
<td>- Invading visceral pleura</td>
</tr>
<tr>
<td></td>
<td>- Atelectasis or obstructive pneumonitis that extends to the hilar region, but does not involve the entire lung</td>
</tr>
<tr>
<td>T2a</td>
<td>Tumor &gt; 3 cm but ≤ 5 cm</td>
</tr>
<tr>
<td>T2b</td>
<td>Tumor &gt; 5 cm but ≤ 7 cm</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor &gt; 7 cm or tumor that invades any of the following:</td>
</tr>
<tr>
<td></td>
<td>- Chest wall, diaphragm, phrenic nerve, mediastinal pleura, parietal pericardium</td>
</tr>
<tr>
<td></td>
<td>- Tumor in the main bronchus &lt; 2 cm distal to the carina, but without involvement of carina</td>
</tr>
<tr>
<td></td>
<td>- Associated atelectasis or obstructive pneumonitis of the entire lung</td>
</tr>
<tr>
<td></td>
<td>- Separate tumor nodule(s) in the same lobe</td>
</tr>
<tr>
<td>T4</td>
<td>Tumor of any size that invades any of the following:</td>
</tr>
<tr>
<td></td>
<td>- Mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophageus, vertebral body, carina</td>
</tr>
<tr>
<td></td>
<td>- Separate tumor nodule(s) in a different ipsilateral lobe</td>
</tr>
</tbody>
</table>

Lymph nodes (N)

<table>
<thead>
<tr>
<th>N0</th>
<th>No regional lymph node metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes, including involvement by direct extension</td>
</tr>
<tr>
<td>N2</td>
<td>Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s)</td>
</tr>
<tr>
<td>N3</td>
<td>Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s)</td>
</tr>
</tbody>
</table>

Distant metastasis (M)

<table>
<thead>
<tr>
<th>M0</th>
<th>No distant metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Distant metastasis</td>
</tr>
<tr>
<td>M1a</td>
<td>Separate tumor nodule(s) in a contralateral lobe; tumor with pleural nodules or malignant pleural (or pericardial) effusion</td>
</tr>
<tr>
<td>M1b</td>
<td>Distant metastasis</td>
</tr>
</tbody>
</table>

By the combination of T, N and M, an overall stage (I-IV) is obtained (Table 3), which gives an estimate of the individual prognosis and basis for therapy [39]. Stage I denote small tumors (≤ 5 cm) with no evidence of invasion,
while stage II tumors are larger and may have metastasized to regional lymph nodes. Stage III includes large tumors that may infiltrate adjacent structures or have lymph node metastasis in the mediastinal or contralateral lymph nodes. Stage IV denotes the most advanced stage of lung cancer, where the tumor has developed distant metastasis [40].

<table>
<thead>
<tr>
<th>Stage I</th>
<th>IA</th>
<th>IB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1a N0 M0</td>
<td>T2a N0 M0</td>
</tr>
<tr>
<td></td>
<td>T1b N0 M0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage II</th>
<th>IIA</th>
<th>IIB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1a N1 M0</td>
<td>T2b N1 M0</td>
</tr>
<tr>
<td></td>
<td>T1b N1 M0</td>
<td>T3 N0 M0</td>
</tr>
<tr>
<td></td>
<td>T2a N1 M0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T2b N0 M0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage III</th>
<th>IIIA</th>
<th>IIIB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1 N2 M0</td>
<td>T4 N2 M0</td>
</tr>
<tr>
<td></td>
<td>T2 N2 M0</td>
<td>Any T N3 M0</td>
</tr>
<tr>
<td></td>
<td>T3 N1 M0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T3 N2 M0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T4 N0 M0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T4 N1 M0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage IV</th>
<th>IV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Any T Any N M1a</td>
<td>Any T Any N M1b</td>
</tr>
</tbody>
</table>

**Table 3. Stage grouping. Reproduced from [38].**

Treatment

Approximately 25% of the NSCLC patients in Sweden are diagnosed with localized disease (stage I or stage II) [20]. These patients are eligible for surgery, with the aim of complete resection and potential to cure the disease [41]. However, even with complete resection, many stage I and stage II NSCLC patients relapse (30-55%) [42]. The 5-year survival rate in patients with stage IA and IB is around 50% and 45%, respectively, while stage IIA and stage IIB present 5-year survival rates of around 35% and 25%, respectively [38]. With adjuvant chemotherapy, stage II patients demonstrate a decreased recurrence rate translating into a moderately (~5%) improved five-year overall survival rate [43-44]. Patients with stage I or II disease who are not eligible for surgery, e.g. because of medical comorbidities, are recommended radiotherapy with curative intent [41].
About 10% of the Swedish NSCLC patients present with stage IIIA disease [20]. This stage describes a heterogeneous group, ranging from resectable tumors with limited nodal metastases to unresectable nodal disease. Patients with limited lymph nodal involvement (N1) are considered for primary surgery followed by adjuvant chemotherapy. In contrast, patients with extended lymph node involvement (N2) may receive neoadjuvant chemoradiotherapy prior to surgical resection [4]. The 5-year survival rate in patients with stage IIIA disease is about 19% [38].

The majority of patients (~65%) are diagnosed with unresectable disease (stage IIIB and stage IV), with extended nodal disease or distant metastasis [20]. These tumors are not curable, and only palliative treatment is available. Depending on the performance status and individual tumor extension, stage IIIB-patients receive platinum-based chemoradiotherapy [4], leading to a 5-year survival of approximately 7% [38]. Stage IV patients usually receive platinum-based chemotherapy to improve the symptoms as well as to increase the survival time [4,45]. The 5-year survival rate in stage IV patients is approximately 2% [38]. Recently, for minor subgroups of advanced stage NSCLC patients, new individualized treatment options in the form of so-called targeted therapies have become available.

**Targeted therapies**

The increasing understanding of cancer biology has identified genetic aberrations believed to be important in tumor growth or progression. This has led to the development of targeted therapies against these specific molecular structures [46]. This personalized approach is applicable in small subsets of patients and is less toxic and have fewer side effects than conventional chemotherapy [47].

EGFR-inhibitors were the first class of targeted compounds approved in advanced NSCLC [5]. Mutations in *EGFR* induce intracellular signaling cascades, causing increased cell proliferation and cell survival [48]. Patients harboring activating *EGFR* mutations are likely to respond to treatment with EGFR tyrosine kinase inhibitors (TKIs) [49]. The first EGFR TKIs approved for treatment of advanced NSCLC were gefitinib and erlotinib. With these small molecule inhibitors, a prolonged progression-free survival has been demonstrated compared to standard chemotherapy (e.g. progression-free survival of 10.8 months with gefitinib compared to 5.4 months with chemotherapy, and 9.7 months with erlotinib compared to 5.2 months with chemotherapy) [50-51].

A second genomic aberration recently discovered in NSCLC is rearrangement of the *ALK* gene. A fusion involving *ALK* and most commonly echinoderm microtubule associated protein-like 4 (EML4) results in a constitutively active ALK, affecting cell proliferation and survival [52]. Targeting
of ALK with a small molecule inhibitor, crizotinib, has shown impressive response rates in patients with ALK rearrangement [6], with a median progression-free survival of 7.7 months compared to 3.0 months with chemotherapy alone [53].

Other molecular aberrations being studied as potential future treatment targets in NSCLC are for example ROS1 [54], RET [55] and BRAF [56].

The specific inhibition of molecules involved in tumor angiogenesis has also been extensively investigated in clinical trials. Bevacizumab is a monoclonal antibody inhibiting vascular endothelial growth factor (VEGF) showing benefit in NSCLC patients (median survival time of 12.3 months for patients receiving bevacizumab in combination with chemotherapy compared to 10.3 months for patients receiving chemotherapy alone) and is approved for first-line treatment of advanced non-squamous lung cancer patients [57-58].

Although targeted therapies have shown impressive response rates in subsets of patients with advanced NSCLC, many patients rapidly develop resistance and experience relapse. Tumors with activating EGFR mutation develop secondary mutations (most commonly T790M), leading to decreased affinity for first-generation TKIs [59]. Similarly, patients receiving ALK-inhibitors eventually develop resistance, e.g. through acquired mutations in the ALK kinase domain, leading to impaired inhibitor function [60]. New agents and strategies to overcome resistance have been developed. Second and third generation EGFR- and ALK-inhibitors are demonstrating effect in patients resistant to first-generation inhibitors, but, resistance is eventually developed to these inhibitors as well [61-64].

Immunotherapy

Over the past few years, immunotherapy has demonstrated encouraging results in patients with NSCLC. Instead of targeting the cancer cells, immunotherapy utilizes and enhances the patient’s own immune system to produce a response against cancer cells [7].

Today, the most promising immunomodulating drugs in NSCLC are so-called check-point inhibitors, modulating T-cell activation. The best established example is antibody-based targeting of the programmed cell death-1 (PD-1) axis [7]. A phase III trial using the PD-1 inhibitor nivolumab in patients with squamous cell carcinoma reported an increased median overall survival of 9.2 months compared to 6.0 months with chemotherapy (NCT01642004), leading to the United States Food and Drug Administration (FDA) approval of nivolumab as second-line therapy in previously treated squamous NSCLC patients [65]. Other clinical trials with inhibitors of PD-1 and the corresponding ligand of PD-1 (PD-L1) are ongoing, e.g. NCT02142738, NCT02220894, NCT02031458, NCT02008277.
Anti-inflammatory agents
While the specific immune response against cancer might inhibit tumor progression, unspecific inflammation preceding or following malignant transformation is considered to have tumor-promoting effects [66]. It is well known that chronic inflammation increases the risk of various cancer types [67], indicating that suppressing inflammation may represent a strategy for cancer prevention and therapy [68]. Already in 1983 it was demonstrated that nonsteroidal anti-inflammatory drugs (NSAIDs) reduced the number of colorectal adenomas in patients with familial adenomatous polyposis [69]. NSAIDs have also been used as chemopreventive agents to decrease the risk of colorectal cancers [70-71].

The initial NSAIDs targeted both of the rate-limiting enzymes involved in the prostaglandin synthesis, cyclooxygenase 1 (COX-1) and COX-2, but due to gastrointestinal side-effects caused by inhibition of COX-1, specific COX-2 inhibitors have been developed [72]. COX-2 is the most frequently evaluated anti-cancer anti-inflammation target. High levels of this enzyme have been demonstrated in various cancer forms [73-76], and the expression is reported to be associated with poor patient outcome [76-78]. Although many studies have tried to elucidate the role of COX-2 in lung cancer, the prognostic impact as well as the clinical benefit of therapeutic inhibition of this enzyme is still not clarified [79-80].

Prognostic and predictive biomarkers
In order to provide information regarding the outcome of a cancer disease, prognostic biomarkers are utilized [81]. The currently most important factors to predict survival and to subsequently guide treatment decisions in NSCLC patients are TNM stage together with performance status and age [82-84]. However, patients with comparable tumor stage, age and performance status may experience strikingly different outcomes with regard to long-term survival as well as early local or metastatic recurrence and death. This strongly suggests that additional factors influence the aggressiveness of the tumor and thus determine the prognosis. Therefore, a systematic correlation of molecular alterations in the tumor tissue to clinical outcome may help to identify biomarkers that aid in prognostication. This is of particular importance for early stage patients, who are potentially curable by complete surgical resection, but have a high risk of tumor recurrence [82]. The addition of adjuvant treatment in patients with stage II disease decreases this risk only modestly [43]. Thus, a majority of stage II patients could be spared from unnecessary treatment, while some stage I patients might benefit from adjuvant therapy. Biomarkers that predict recurrence and that are applicable in routine clinical diagnostics are therefore greatly needed [82].
Predictive biomarkers, on the other hand, are used to predict and monitor the response, or absence of response, to a specific treatment [85]. In lung cancer, oncogenic driver mutations present such valuable predictive biomarkers [86], and are often also targets for therapy [87]. Currently, \textit{EGFR} mutations and \textit{ALK} rearrangements are predictive for tumor response in patients treated with the corresponding inhibitor [4,86].

**Immunohistochemistry (IHC)**

Immunohistochemistry (IHC) is a commonly used technique in pathological diagnostics as well as in research. With this technique, antibodies are utilized for visualization of protein expression in tissue sections. When performing IHC, the tissue morphology is preserved and information about the distribution and location of the protein expression within the tissue is obtained [88]. The concept of IHC is based on the interaction between a specific antibody and the precise position (epitope) on its respective protein (antigen) [89].

Antibodies are a group of proteins called immunoglobulins (Ig) that are present in the blood. They are produced by plasma cells in response to invading pathogens in the body [90]. Immunoglobulins can be divided into five major classes: immunoglobulin G (IgG), IgA, IgD, IgE and IgM. IgG is the most abundant immunoglobulin in the serum, and is therefore most commonly used in the production of antibodies for immunoassays. In immunohistochemical reactions, two types of antibodies are used; monoclonal and polyclonal. Monoclonal antibodies are created by the same plasma cell clone, which make all the antibodies identical and directed against a single epitope on an antigen [91]. Polyclonal antibodies are produced by several plasma cell clones and comprise a heterogeneous mixture of antibodies directed against multiple epitopes on the same antigen [91-92].

There are two ways of visualizing the immune reaction, directly or indirectly (Figure 1A and B). With the direct method, the antigen is detected by the binding of a primary antibody that is directly labeled with an enzyme, biotin or fluorophore [89,91]. However, the most common way to visualize the immune reaction is by the indirect method. With this method an unlabeled primary antibody reacts with the antigen in the tissue, followed by the binding of a labeled secondary antibody to the primary antibody. The secondary antibody can be labeled with different enzymes, e.g. alkaline phosphatase (AP) or horseradish peroxidase (HRP) [89]. HRP is most commonly used, and when adding hydrogen peroxide (H$_2$O$_2$) and the chromogen 3,3’-diaminobenzidine (DAB), a brown staining is produced where the primary and secondary antibodies are attached. This is known as DAB staining [93]. For low density antigens, it is recommended to employ further amplification
of the IHC signal. Amplification of a signal occurs when more label is deposited per primary antibody compared to when using a standard primary-secondary antibody arrangement. Amplification can be accomplished by for instance the avidin-biotin complex, a chain polymer-conjugated technology such as the “EnVision” system developed by Dako-Cytomation, the “OptiView” hapten-based system from Ventana Medical Systems or with tyramide signal amplification. With the avidin-biotin complex, the secondary antibody is labeled with biotin. Streptavidin (or avidin) labeled with HRP is added and binds to biotin on the secondary antibodies, and when adding a chromogen, e.g. DAB, a brown staining is produced (the ABC method) [91,94]. The chain polymer-conjugated technology uses an enzyme-labeled dextran backbone to which several enzyme-molecules and secondary antibodies are attached. As many as 100 enzyme molecules (HRP or AP) and up to 20 secondary antibodies can be linked to a single backbone, and when a chromogenic substrate and H₂O₂ are added, a several-fold higher antigen detection is obtained compared to when using standard indirect protocols [95] (Figure 1C). The hapten-based technology from Ventana also uses a similar principle - the secondary antibody is conjugated to several non-endogenous 3-hydroxy-2-quinoxaline (HQ) haptens [96]. A hapten is a small chemical that is not antigenic in itself, but has to be conjugated to an endogenous or exogenous protein, usually an antibody, to become antigenic [97]. To the haptns, tertiary antibodies linked to several HRP enzymes binds, and when adding DAB and H₂O₂, a brown staining is produced [96] (Figure 1D). Another way of amplifying the IHC-signal is by tyramide signal amplification. In this method, HRP-labeled secondary antibodies or HRP-avidin complexes are used. When H₂O₂ and labeled tyramine is added to the secondary antibodies, the tyramine is oxidized by HRP, leading to the binding of tyramine molecules in the immediate surrounding of the immunoreaction. The tyramine can be labeled with a fluorochrome, which can be visualized in a fluorescence microscope, or with biotin, which is visualized via labeled streptavidin or avidin in a light microscope [91,98]. To amplify the signal even further, the OptiView technology offers an optional step based on the tyramide signal amplification system. After the binding of secondary antibodies with conjugated HQ-haptens, a HQ-tyramide complex can be added. This complex reacts with H₂O₂ and HRP on the tertiary antibodies, leading to the binding of tyramine molecules in the surrounding of the immunoreaction. This is followed by the addition of HRP-linked quaternary antibodies binding to the haptns in the complex, and when adding DAB, a highly amplified signal is obtained [96,99].
Polymer backbone
HRP
HQ-Hapten
Tertiary antibody
Secondary antibody
Primary antibody

Tissue sample with antigen

A) Direct detection
B) Indirect detection
C) Polymer-based technology e.g. “EnVision”
D) Hapten-based technology e.g. “OptiView”

Figure 1. The principle of immunohistochemical staining with detection by the direct method (A) and by the indirect method (B). Further amplification of the IHC signal can be obtained by using the polymer-based technology here exemplified by “EnVision” (C) or with the hapten-based technology, such as “OptiView” (D).

Fluorescence in situ hybridization (FISH)

Another way of studying biological events in their morphological localization is by in situ hybridization (ISH). ISH is used for visualization of nucleic acids sequences in their “original place” e.g. in their cellular context [100]. The technique was first reported in 1969 [101], then with the use of radio-labeled probes, but was later further developed and the probes were instead labeled with fluorescent dye (fluorochromes), hence the name fluorescence in situ hybridization (FISH) [102]. One of the first applications of FISH was chromosome painting, where entire chromosomes are visualized [103-104], but later single-locus probes were developed making it possible to detect chromosomal aberrations such as deletions, duplications or rearrangements of genetic material [104-105].

For the detection of a specific genomic region, a probe that is complementary to a part of that region is used. To visualize the probes, they have to be labeled. The labeling can be performed in two ways, directly or indirectly. When labeled directly, the probes are conjugated to a fluorochrome, e.g.
Texas red or fluorescein, while with indirect labeling the molecule attached to the probe is non-fluorescent, for example either biotin or a hapten, and then incubated with avidin or an antibody that is fluorescently labeled [102,106]. Many probes, directly or indirectly labeled, are commercially available. The probes are applied on the tissue section and after denaturation the probes hybridize to the complementary target DNA [102]. If the probe is fluorescently labeled, the sample is directly counterstained with 4',6-diamidino-2-phenylindole (DAPI), which is a dye binding to AT-rich regions in the DNA and commonly used for nuclear counterstaining [91]. The signals are then observed with a fluorescent microscope equipped with filters that visualize the different wavelengths of the fluorescent probes. If the labeling of the probe is indirect, the addition of a fluorescently labeled antibody is necessary before the sample can be counterstained with DAPI and the probes visualized [102]. The principle of direct and indirect FISH is outlined in Figure 2.
In order to detect gene rearrangements, two different types of probes can be utilized: fusion probes and break-apart probes. With fusion probes, each of the two genes involved in the translocation are detected with differently labeled probes (e.g. red and green). When no rearrangement has occurred, two separate color signals are visualized. However, if rearranged, the differently labeled probes will be very close or on top of each other and only one color will be observed (yellow) [107]. With break-apart probes, differently labeled probes (e.g. red and green) are hybridized on each side of the break point on one of the genes involved in the rearrangement. If no rearrangement has occurred, the probes will be very narrow or on top each other (yellow signal), but if the sample is rearranged, there will be a split between the signals and the green and red signals will be separated. With fusion probes only known genetic partners can be detected, while with break-apart probes, the
fusion partner remains unidentified [108]. The principle of fusion and break-apart probes is illustrated in Figure 3.

**FISH fusion probes**

![FISH fusion probes diagram]

**FISH break-apart probes**

![FISH break-apart probes diagram]

*Figure 3.* The principle of FISH fusion probes and break-apart probes. With fusion probes, each of the two genes involved in the translocation are detected with differently labeled probes (e.g. red and green). If the sample is not rearranged, two separate color signals are visualized. If the sample is rearranged, the differently labeled probes will be very close or on top of each other and only one color will be observed (yellow). With break-apart probes, differently labeled probes (e.g. red and green) are hybridized on each side of the break point on one of the genes involved in the rearrangement. If the sample is not rearranged, the probes will be very narrow or on top each other (yellow signal), but if the sample is rearranged, a split between the signals will be displayed (the red and green signal will be separated).

**Tissue microarray (TMA)**

At pathological laboratories patient tissue is most often available in the form of formalin fixed paraffin embedded (FFPE) material. FFPE material is a valuable resource, since fixation in formalin and embedding in paraffin preserves the protein and the DNA as well as the tissue morphology. With paraffin embedding, the samples are also easy to store [109-110]. Sections from the FFPE blocks can then be used in different analyses. However, to save
reagents and tissue, a tissue microarray (TMA) can be utilized. A TMA is composed of cylindrical cores from many different primary FFPE tumor blocks collected in a single “recipient” paraffin block. Therefore, it is possible to simultaneously examine a large number of tissues on a single slide [111-113]. In order to collect representative tissue cores, the primary tissue blocks are cut in 4 μm sections, placed on glass slides, and stained with hematoxylin and eosin (HE) to visualize the morphology of the tissue. Representative areas are then marked on the glass slide, and cores from that area are punched from the tumor block using a manual or automated arraying instrument and deposited in the recipient block [113]. The generation of a TMA is presented in Figure 4.

![Figure 4. Generation of a tissue microarray (TMA). Hematoxylin and eosin (HE) stained sections from primary tissue blocks are marked for representative tissue areas. Tissue cores from these primary blocks are then collected in a recipient block by the use of an arraying instrument. The generated TMA-block is sectioned, and the sections are placed on glass slides and used in immunohistochemical analyses.](image)

By assembling tissue cores from primary donor blocks into a TMA, the majority of the tissue in primary donor blocks is saved even though new analyses are performed on the patient sample [112]. Up to 300 sections can be cut from a single TMA block, and the sections can be utilized for several different in situ analyses, for example IHC, mRNA in situ hybridization and FISH [112-113]. TMAs containing tumor specimens are particularly useful for linking protein expression from various markers to clinical parameters in order to screen for their prognostic, predictive or diagnostic significance in patients with a certain cancer disease [114].

**Gene expression microarray**

Gene expression microarrays make it possible to simultaneously analyze mRNA expression levels of thousands of genes from a biological sample at a certain time point [115-116]. A probe on an Affymetrix Human Genome U133 Plus 2.0 array consists of a sequence of 25 oligonucleotides, and each
probe is synthesized on the surface of a chip. Commonly, 11 different perfect-match probes, which bind to different positions on a single mRNA sequence, constitute a probe set [117]. The Human Genome U133 Plus 2.0 array contains more than 54,000 probe sets, representing over 47,000 transcripts [118]. In order to perform a microarray analysis, RNA is first extracted from a sample and converted to complementary DNA (cDNA) by the use of reverse transcriptase. The cDNA is then converted to complementary RNA (cRNA) by an \textit{in vitro} transcription reaction where biotinylated nucleotides are incorporated. The cRNA is then added to a glass chip and hybridized to the complementary gene sequences present on the chip. Non-bound material is washed away. The fluorescence marker streptavidin-phycoerythrine, which binds to biotin, is then added to the samples, the chip is scanned, and the signals emitted are detected [119]. This procedure is visualized in Figure 5.

\textbf{Figure 5.} Gene expression microarray analysis using Affymetrix GeneChip. RNA is extracted from frozen tissue samples and converted to cDNA, followed by conversion to cRNA with subsequent incorporation of biotinylated nucleotides. The cRNA is then fragmented and hybridized to the Affymetrix GeneChip. Signals from fluorescent markers which have bound to biotin are detected and analyzed.
Present Investigation

Aim
The main objective of this thesis was to evaluate the expression of possible new treatment targets and of biomarkers that can guide therapy in NSCLC. This was done on a multi-molecular level integrating genomic, gene expression and protein data with clinicopathological parameters.

The specific aims for each study were:

Paper I: To evaluate expression levels and survival associations of the potential treatment targets claudin 6 and claudin 18.2 in NSCLC.

Paper II: To evaluate the prognostic influence of tumor and stromal protein expression of the inflammatory mediator cyclooxygenase 2 (COX-2) in two independent NSCLC cohorts.

Paper III: To determine if elevated tumor and stromal protein expression of COX-2 is associated with improved survival in patients treated with the selective COX-2 inhibitor celecoxib.

Paper IV: To compare the results of two different ALK detection methods, FISH and IHC, and to analyze the overall frequency of ALK rearrangements in a large representative Swedish NSCLC population.

Molecular targets studied
Claudin 6 and claudin 18.2
Claudins are essential components of tight junctions and are responsible for the regulation of paracellular activity between cells [120-121]. Individual cells express more than one of the 26 human claudin family members [122], suggesting that each tissue has its own combination of claudin proteins and regulation of tight junction function [123-124]. Many of the claudins are broadly expressed, while others such as claudin 6 and 18.2 are selectively expressed. The expression of claudins has been demonstrated to be altered in
several tumor types [124]. Claudin 6 expression is normally limited to the embryonic development, but it is expressed in various cancer types [124-125]. Claudin 18 has two isoforms, claudin 18.1, which is constitutively expressed in lung tissue [126], and claudin 18.2, which is expressed only in differentiated epithelial cells of the gastric mucosa. Claudin 18.2 has also been demonstrated to be ectopically expressed in various cancer types [127].

Cyclooxygenase 2 (COX-2)

Cyclooxygenase 2 (COX-2) is the most frequently studied anti-cancer anti-inflammatory target today [68]. COX-2 is encoded by the gene prostaglandin-endoperoxide synthase 2 (PTGS2), and is the rate-limiting enzyme responsible for the conversion of free arachidonic acid to prostaglandins, prostacyclins and thromboxanes [128-129]. COX-2 is induced by growth factors, cytokines, tumor promoters, oncogenes and other extracellular stimuli [130], and subsequently, high levels of COX-2 have been demonstrated in various malignancies [131-134]. Due to its pro-tumorigenic properties, COX-2 has been evaluated in the context of cancer prevention and treatment [135-137], and based on the promising results of COX-2 inhibition in colorectal cancer [138], COX-2 inhibitors have also been evaluated for the treatment of NSCLC patients [139].

Anaplastic lymphoma kinase (ALK)

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase that belongs to the insulin growth factor receptor superfamily [140]. The function of ALK is still unknown, but ALK is believed to play a role in the development and maintenance of the nervous system since the protein normally is expressed only in a small number of cells within the developing and mature nervous system [140-141]. Rearrangements of ALK are found in 3-7% of NSCLC patients [26], and this rearrangement leads to a constitutive active ALK, affecting downstream signaling and promoting malignant growth and proliferation. Most commonly, ALK forms a fusion protein with EML4 [52], but several other fusion partners have been described. All fusion proteins contain the ALK kinase domain, subsequently leading to comparable kinase activation [52,142-144]. ALK rearrangements can be targeted by small synthetic ALK-inhibitors, leading to impressive tumor shrinkage and improved survival in patients harboring this rearrangement [6,53].
Material and Methods

Patient cohorts and study design

In this thesis, four NSCLC patient cohorts were studied: the Uppsala I cohort (Paper I, II and IV), the Uppsala II cohort (Paper IV), the Örebro cohort (Paper II and IV) and the Cyclus cohort (Paper III).

The Uppsala I cohort consists of operated patients with primary NSCLC reported to the Uppsala-Örebro Regional Lung Cancer Registry between 1995 and 2005, with available fresh frozen tissue in the Uppsala frozen tissue Biobank. When the Regional Lung Cancer Registry was cross-linked to the frozen tissue Biobank database, 382 patients with corresponding frozen tissue blocks were identified. Study inclusion was based on a) confirmed NSCLC histology in the fresh frozen section, b) a tumor sample size of ≥ 5mm, and c) a tumor cell fraction of ≥ 50%. Patients who had received neoadjuvant treatment were excluded. In total, 194 patients fulfilled the criteria. From these samples RNA was previously extracted and analyzed by Affymetrix gene expression microarrays [145-146]. Additionally, FFPE tissue was available for 189 of these patients plus 165 additional tumor samples (where the corresponding frozen tissue specimen did not meet the previously mentioned inclusion criteria), ending up with in total 354 patients with available FFPE tissue. Duplicate tissue cores from each patient sample were previously included in a TMA [147-148]. Information about histology was attained from the pathology reports at Uppsala University Hospital. Clinical patient characteristics, e.g. performance status according to WHO, stage and smoking history, was obtained from the records of the Uppsala-Örebro Regional Lung Cancer Registry.

The Uppsala II cohort is an extension of the Uppsala I cohort, consisting of patients with primary NSCLC that had undergone surgical resection at the Uppsala University Hospital between 2006 and 2010. When the Regional Lung Cancer Registry was crosslinked to the Laboratory Information System (Sympathy), a total number of 390 patients were identified. After confirming histopathological examination and exclusion of patients who had received neoadjuvant treatment, 354 samples remained. Tumor cores from these FFPE samples were included in a TMA, with duplicate tissue cores from each patient. Clinical patient data were obtained from the records of the Uppsala-Örebro Regional Lung Cancer Registry [33,149].

The Örebro cohort consists of 262 patients that had undergone surgical resection at the Örebro University Hospital between 1990 and 1995. A TMA comprising triplicate cores from each tumor specimen was available. Clinical data, including histology, stage, age, smoking status and overall survival had previously been obtained from the clinical, surgical and pathological records at the Örebro University Hospital [150-151].
The study design of the Uppsala I cohort, the Uppsala II cohort and the Örebro cohort is described and visualized in Figure 6. Clinical characteristics of patients included in the three cohorts are presented in Table 4. In the statistical analyses, overall survival was computed from the date of diagnosis to the date of death or last follow-up until spring 2011 (Uppsala I cohort), summer 2015 (Uppsala II cohort) or autumn 2010 (The Örebro cohort).

Figure 6. Study design of the cohorts. (A) The Uppsala I cohort consists of NSCLC patients that had undergone surgical resection between 1995 and 2005. FFPE tissue samples from 354 patients were utilized for protein expression analysis using IHC on TMAs, while 194 fresh frozen tissue samples were subjected to gene expression analysis. (B) The Uppsala II cohort consists of NSCLC patients that were operated between 2006 and 2010. FFPE tissue samples from 354 patients were subjected to protein expression analysis using IHC on TMAs. (C) The Örebro cohort consists of NSCLC patients that underwent surgical resection between 1990 and 1995. FFPE tissue samples from 262 patients were subjected to protein expression analysis using IHC on TMAs.
Table 4. Clinical characteristics of the lung cancer patients included in the gene expression patient cohort (Uppsala I cohort), the Uppsala I TMA cohort, the Uppsala II TMA cohort as well as the Örebro TMA cohort.

<table>
<thead>
<tr>
<th></th>
<th>Uppsala I TMA</th>
<th>Uppsala I Affymetrix</th>
<th>Uppsala II TMA</th>
<th>Örebro TMA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IHC, n (%)</td>
<td>GE, n (%)</td>
<td>IHC, n (%)</td>
<td>IHC, n (%)</td>
</tr>
<tr>
<td>All patients</td>
<td>354 (100)</td>
<td>194 (100)</td>
<td>354 (100)</td>
<td>262 (100)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>188 (53.1)</td>
<td>105 (54.1)</td>
<td>173 (48.9)</td>
<td>165 (63.0)</td>
</tr>
<tr>
<td>Female</td>
<td>166 (46.9)</td>
<td>89 (45.9)</td>
<td>181 (51.1)</td>
<td>97 (37.0)</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 70 years</td>
<td>244 (68.9)</td>
<td>150 (77.3)</td>
<td>232 (65.5)</td>
<td>189 (72.1)</td>
</tr>
<tr>
<td>&gt; 70 years</td>
<td>110 (31.1)</td>
<td>44 (22.7)</td>
<td>122 (34.5)</td>
<td>73 (27.9)</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current + former smoker</td>
<td>318 (89.8)</td>
<td>179 (92.3)</td>
<td>313 (88.4)</td>
<td>206 (78.6)</td>
</tr>
<tr>
<td>Never smoker</td>
<td>34 (9.6)</td>
<td>15 (7.7)</td>
<td>41 (11.6)</td>
<td>19 (7.3)</td>
</tr>
<tr>
<td>Missing data</td>
<td>2 (0.6)</td>
<td>0</td>
<td>0</td>
<td>37 (14.1)</td>
</tr>
<tr>
<td>TNM-stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>87 (24.6)</td>
<td>38 (19.6)</td>
<td>146 (41.2)</td>
<td>86 (32.8)</td>
</tr>
<tr>
<td>IB</td>
<td>150 (42.4)</td>
<td>90 (46.4)</td>
<td>76 (21.5)</td>
<td>33 (12.6)</td>
</tr>
<tr>
<td>IIA</td>
<td>12 (3.4)</td>
<td>6 (3.1)</td>
<td>41 (11.6)</td>
<td>34 (13.0)</td>
</tr>
<tr>
<td>IIIB</td>
<td>44 (12.4)</td>
<td>29 (14.9)</td>
<td>34 (9.6)</td>
<td>50 (19.1)</td>
</tr>
<tr>
<td>IIIA</td>
<td>35 (9.9)</td>
<td>21 (10.8)</td>
<td>50 (14.1)</td>
<td>43 (16.4)</td>
</tr>
<tr>
<td>IIIB</td>
<td>16 (4.5)</td>
<td>6 (3.1)</td>
<td>0 (0.0)</td>
<td>2 (0.8)</td>
</tr>
<tr>
<td>IV</td>
<td>10 (2.8)</td>
<td>4 (2.1)</td>
<td>7 (2.0)</td>
<td>2 (0.8)</td>
</tr>
<tr>
<td>Missing data</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12 (4.6)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>195 (55.1)</td>
<td>106 (54.6)</td>
<td>210 (59.3)</td>
<td>134 (51.1)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>117 (33.1)</td>
<td>63 (32.5)</td>
<td>104 (29.4)</td>
<td>124 (47.3)</td>
</tr>
<tr>
<td>Large cell cancer/NOS(^1)</td>
<td>42 (11.9)</td>
<td>25 (12.9)</td>
<td>40 (11.3)</td>
<td>4 (1.5)</td>
</tr>
<tr>
<td>Performance status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>187 (52.8)</td>
<td>105 (54.1)</td>
<td>214 (60.5)</td>
<td>n/a(^2)</td>
</tr>
<tr>
<td>1</td>
<td>134 (37.9)</td>
<td>73 (37.6)</td>
<td>137 (38.7)</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>27 (7.6)</td>
<td>12 (6.2)</td>
<td>3 (0.8)</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>5 (1.4)</td>
<td>4 (2.1)</td>
<td>0 (0.0)</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>1 (0.3)</td>
<td>0</td>
<td>0</td>
<td>n/a</td>
</tr>
<tr>
<td>Follow-up mean, months</td>
<td>59.9</td>
<td>58.4</td>
<td>49.9</td>
<td>49.3</td>
</tr>
</tbody>
</table>

\(^1\)NOS = not otherwise specified
\(^2\)n/a = not applicable

The Cyclus cohort comprises tissue from NSCLC patients from a Swedish randomized phase III trial [152]. The trial was designed to investigate if the addition of the COX-2 inhibitor celecoxib to standard chemotherapy would prolong survival in patients with advanced NSCLC. Chemotherapy-naïve patients with advanced (stage IIIB and IV) NSCLC were included in the trial (n=316). From the original trial population of 316 patients, biopsies and surgical specimens from 122 (38.6%) patients prior to the administration of celecoxib were available for protein expression analysis. The Cyclus study
design is described and visualized in Figure 7. Clinical characteristics of the patients with available tissue for protein analysis are presented in Table 5. Overall survival was computed from the date of diagnosis to the date of death or last follow-up until summer 2015.

Table 5. Clinical characteristics of the 122 lung cancer patients with available tissue specimens included in the Cyclus trial.

<table>
<thead>
<tr>
<th></th>
<th>Placebo (n=55)</th>
<th>Celecoxib (n=67)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (range)</td>
<td>64 (47-84)</td>
<td>66 (38-85)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>31 (56%)</td>
<td>31 (46%)</td>
</tr>
<tr>
<td>Female</td>
<td>24 (44%)</td>
<td>36 (54%)</td>
</tr>
<tr>
<td>Performance status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>14 (25%)</td>
<td>16 (24%)</td>
</tr>
<tr>
<td>1</td>
<td>32 (58%)</td>
<td>37 (55%)</td>
</tr>
<tr>
<td>2</td>
<td>9 (16%)</td>
<td>14 (21%)</td>
</tr>
<tr>
<td>TNM-stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIIB</td>
<td>15 (27%)</td>
<td>21 (31%)</td>
</tr>
<tr>
<td>IV</td>
<td>40 (73%)</td>
<td>46 (69%)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>32 (58%)</td>
<td>42 (63%)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>18 (33%)</td>
<td>21 (31%)</td>
</tr>
<tr>
<td>Large cell carcinoma/NOS*</td>
<td>5 (9%)</td>
<td>4 (6%)</td>
</tr>
</tbody>
</table>

*NOS = not otherwise specified
Methodological considerations

Immunohistochemistry
Immunohistochemistry (IHC) was used to evaluate protein expression in all Papers (I-IV) included in this thesis. The staining was performed on FFPE slides using an automated staining instrument, making the staining more reproducible compared to when manual staining is performed. The IHC-staining was manually evaluated on scanned slides using a semi-quantitative annotation method where staining intensity and fraction of cells stained were taken into consideration. The intensity and fraction of stained cells were assigned ordinal scores, which were multiplied to obtain an overall staining score for each sample. The overall score was used in the subsequent analysis. As the evaluation of the staining is subjective, we attempted to reduce this uncertainty by the independent evaluation by two observers. In case of any discordancy, a shared review of the images was performed to obtain a common interpretation.

Since staining results can be influenced by many factors, we also compared two staining protocols using the same antibody clone, but differing with regard to staining instrument, pre-treatment method, incubation time of primary antibody and signal amplification method (Paper IV).

One main advantage of IHC is that protein expression can be allocated to a cell type or a tissue compartment. This was used to determine compartment-specific expression in Paper II and Paper III, where we analyzed protein expression in tumor cells and stroma cells separately.

Tissue microarrays
In Paper I, II and IV, the protein expression was evaluated on TMAs. One potential problem with TMAs is that the included tissue cores may not be representative for the whole tumor. However, many studies have shown that two or three cores of the original tissue compensate for this heterogeneity and give reliable results [114,153-154]. To address this aspect, in Paper IV we also analyzed the corresponding whole section with IHC and FISH to directly compare the results. In Paper III, IHC was on the other hand performed on whole sections since no TMA was available for the patients included in this study.

FISH
In Paper IV, ALK rearrangements were detected using FISH break-apart probes on TMAs. FISH has the advantage to identify ALK rearrangements independent of fusion partner or break-point. In general, interpretation of
FISH results is considered to be clear-cut, but the ALK specific inversion of a small region on chromosome 2 leads to narrow split signals that can be hard to detect. In accordance with the guidelines published by the International Association for the Study of Lung Cancer (IASLC) [155], we considered a sample ALK FISH positive if ≥ 15% of a minimum of 50 cells evaluated are rearranged.

**Gene expression analysis**

Gene expression data, generated using Affymetrix GeneChip HG-U133 Plus 2.0 arrays, was included as a supplement to IHC (Paper I, II and IV) and FISH (Paper IV) for a subset of patients in the Uppsala I cohort. However, it should here be noted that RNA was extracted from another part of the same tumor, which had been snap frozen and stored at -80ºC, and that one cannot exclude that intratumoral differences might influence the comparison. Furthermore, the gene expression analysis was performed based on RNA extracted from crude tumor tissue and for each gene the signal thus represents the average gene expression of all cell types present. Consequently, gene expression levels are difficult to compare to compartment-specific protein expression levels determined by IHC.

**Statistics**

The Kaplan-Meier method was used to calculate the survival rates. Survival curves for analyzed parameters were compared using the log-rank test. Uni- and multivariate Cox-regression analysis, with inclusion of clinicopathological parameters (age, stage, performance status), was performed to estimate hazard ratios (HR). Protein expression was dichotomized into high or low expression in each paper. Gene expression values were dichotomized into high or low expression by outlier sum statistics [156] in Paper I, and used as a continuous variable in Paper II and IV. Gene expression and protein expression values were compared using Spearman correlation (Paper I) or the Wilcoxon test (Paper II). Gene and/or protein expression and clinical parameters were compared using Spearman correlation (Paper I, II, IV). In Paper II, meta-analyses (random effect and fixed effects models) were conducted for combining single estimates from previous independent studies into one combined overall estimate. All p-values were two-sided and a statistical significance level of p < 0.05 was used. The statistical analyses were performed using the statistical computing language R (https://www.r-project.org) (Paper I, II and IV) and the program Stata (version 13.1) (Paper III).
Results and discussion

Paper I:

*Claudin 6 and claudin 18.2 are aberrantly expressed in subgroups of non-small cell lung cancer.*

Lineage-specific claudins, which are normally confined to a certain cell-type, have been proposed as therapy targets as they may be ectopically expressed in neoplastic tissue and are not expressed in most normal tissues [124]. Clinical trials with therapeutic antibodies targeting claudin 6 in ovarian cancer and claudin 18.2 in gastro-esophageal cancer are currently ongoing (NCT02054351 and NCT01630083). To evaluate if claudin 6 and claudin 18.2 are expressed in NSCLC, and as such, could represent potential treatment targets also in this cancer type, protein expression was analyzed by immunohistochemistry on tissue microarrays (Uppsala I cohort).

Membranous expression of claudin 6 and claudin 18.2 was observed in 6.5% and 3.7%, respectively. Expression of claudin 6 and claudin 18.2 was observed in principally non-overlapping subgroups of non-squamous carcinoma, while non-malignant bronchial epithelial cells and pneumocytes were negative. High expression of claudin 6 was associated with shorter overall survival in the adenocarcinoma subgroup in both the univariate (HR=1.8, CI 1.1-2.9, p=0.032) and multivariate Cox-regression model (HR=1.9, CI 1.2-3.2, p=0.017). High expression of claudin 18.2 was primarily observed in patients with early-stage disease and in tumors with a low Ki67-index, suggesting a less aggressive phenotype. However, a longer survival was not demonstrated in the Cox-regression model. Notably, none of the samples with claudin 6- or claudin 18.2-positivity demonstrated an *EGFR* mutation, which may indicate that claudin 6 and 18.2 expressing patients constitutes a subgroup that is not overlapping with one of the most important markers guiding treatment decisions in advanced NSCLC today.

Our protein findings were supplemented with gene expression microarray-based mRNA data for claudin 6 and total claudin 18. Protein and gene expression could be compared in a subset of 189 samples, and strong correlations were observed for both claudin 6 (rho 0.6-0.8) and claudin 18 (rho 0.4-0.6). However, the gene expression data did not distinguish between the two claudin 18 splice variants, and should therefore be interpreted with caution. High expression of claudin 6 was not associated with survival in the univariate analysis, but one of the analyzed probe sets (208474_at) demonstrated in the multivariate analysis an association with shorter survival (HR=1.6, CI 1.0-2.4, p=0.04). This association was more prominent in the
adenocarcinoma subgroup. For claudin 18, no association with survival was observed. Thus, the gene expression data supports the protein expression data.

In conclusion, claudin 6 and 18.2 are both ectopically expressed in a subset of NSCLC patients. This finding might be of clinical relevance, as these claudins represent cancer specific targets and initial clinical trials with therapeutic antibodies against both claudins may be extended to NSCLC. However, due to the low number of positive cases, further analyses are needed in a larger sample set to evaluate the expression pattern of claudin 6 and 18.2 and if they are confined to a certain subgroup of NSCLC.

Paper II:

*COX-2 is expressed in both tumor and stromal cells in NSCLC. COX-2 protein expression in tumor cells is not associated with survival, while the prognostic significance of stromal COX-2 expression remains unclear.*

COX-2 is an enzyme which is induced and expressed in response to growth factors, cytokines and other stimuli [130]. Strategies to target COX-2 with anti-inflammatory drugs have shown promising results in colorectal cancer prevention [157] and have been demonstrated to reduce the number of colorectal adenomas in patients with familial adenomatous polyposis [158]. However, the role of COX-2 in NSCLC is controversial, with contradicting results regarding its prognostic value. We therefore evaluated the COX-2 gene (*PTGS2*) and protein expression in NSCLC tissue. *PTGS2* expression based on Affymetrix gene expression microarray data was evaluated in 194 patients (Uppsala I cohort). Additionally, a meta-analysis was performed, including nine publically available data sets, including in total 1337 patients. By immunohistochemical analysis, the *in-situ* COX-2 protein expression pattern was evaluated in a compartment-specific manner, distinguishing between tumor cell positivity and positivity in the tumor stroma. The IHC analysis was performed on tissue microarrays from two independent cohorts (Uppsala I and Örebro).

In the Affymetrix gene expression microarray analysis, *PTGS2* expression did not show any significant statistical association with overall survival or other investigated clinical parameters (all p > 0.1). Similarly, no association with survival was demonstrated in the meta-analysis comprising nine publically available data sets, either with all histologies included (p=0.57) or when the histological subgroups were analyzed separately. Hence, no prognostic relevance of *PTGS2* expression could be demonstrated.
In the second phase, we evaluated immunohistochemical COX-2 protein expression in the tumor cell and the stromal cell compartment. Tumor COX-2 expression did not demonstrate any significant correlations with clinical parameters or overall survival in the Uppsala I (HR=0.91, CI 0.69-1.2, adj. p=1.0) or the Örebro cohort (HR=1.18, CI 0.81-1.72, adj. p=1.0). COX-2 expression in stromal cells did also not correlate with any clinical parameters or with overall survival in the Uppsala I cohort (HR=0.86, CI 0.64-1.17, adj. p=1.0). However, in the Örebro cohort, a significant association was found between high stromal COX-2 expression and longer survival. This was observed in both uni- (HR=0.56, CI 0.38-0.82, adj. p=0.018) and multivariate Cox-regression analysis (HR=0.56, CI 0.38-0.84, adj. p=0.02), and was more profound in the adenocarcinoma subgroup (HR=0.42, CI 0.23-0.75, adj. p=0.02).

In summary, the present findings dismiss the prognostic relevance of \textit{PTGS2} expression and reject the hypothesis of COX-2 tumor cell expression being of prognostic value in NSCLC. On the other hand, COX-2 protein expression in the stromal compartment was associated with longer survival in one of the analyzed cohorts, which makes COX-2 stromal expression an interesting subject for further investigation.

\textbf{Paper III:}

\textit{The addition of celecoxib to standard chemotherapy did not prolong overall survival in advanced NSCLC patients, regardless of COX-2 expression.}

Clinical trials have failed to demonstrate a benefit of adding COX-2 inhibitors to standard chemotherapy in NSCLC [159-161]. In Paper III, we aimed to analyze COX-2 expression in tumor and stromal cells as a predictive biomarker for longer survival upon COX-2 inhibition.

Between 2006 and 2009, 316 NSCLC patients were enrolled in a randomized, double-blind, placebo-controlled, multicenter phase III trial (the Cyclus trial) with the primary objective to compare overall survival between two treatment arms: arm one (n=158) received the COX-2 inhibitor celecoxib in addition to standard chemotherapy and arm two (n=158) received placebo and standard chemotherapy. No survival difference was found between the two treatment arms [152]. However, COX-2 expression was not considered as a marker for celecoxib treatment. Therefore, we evaluated COX-2 expression as a biomarker for survival benefit in 122 of the patients included in the Cyclus trial.
In the first step, an updated survival analysis of the patients enrolled in the original Cyclus trial (n=316) as well as of the patients with available tissue for COX-2 analysis (n=122) was performed. No significant survival difference was seen between the two treatment arms, neither in the complete population (p=0.93) nor in the subpopulation available for COX-2 analysis (p=0.52).

In the second step, we analyzed patient survival according to high and low COX-2 protein expression in tumor cells without taking the celecoxib treatment into account, however, no significant survival difference was observed (HR=1.07, CI 0.74-1.54, p=0.74). When comparing high and low COX-2 expression in stromal cells, there was a slight tendency toward a longer survival in patients with positive COX-2 staining, but the difference was statistically non-significant (HR 0.80, CI 0.56-1.15; p=0.23).

In the third step, COX-2 expression was analyzed in accordance to treatment arm. Of the 122 patients available for COX-2 protein expression analysis, 67 received celecoxib in addition to standard chemotherapy. The overall survival difference between patients with high and low COX-2 expression in tumor cells in this group was not significant (HR=0.91, CI 0.55-1.50, p=0.72). As expected, also in the group of patients who received placebo (n=55) no survival difference was seen between patients with high and low COX-2 expression (HR=1.30, CI 0.74-2.26, p=0.36).

A similar pattern was seen in analysis of the COX-2 stromal expression. The survival did not differ between high and low stromal expression in patients receiving celecoxib (HR=0.97 for high vs. low expression, CI 0.59-1.58, p=0.89). In patients receiving placebo, there was a trend towards longer survival time in patients with high COX-2 expression in stroma, although the difference did not reach statistical significance (HR=0.63, CI 0.36-1.10; p=0.10).

In conclusion, the addition of celecoxib to standard chemotherapy did not prolong overall survival in advanced NSCLC patients, irrespective of COX-2 protein expression. However, the retrospective evaluation of COX-2 expression had to be made on available tissue, including 122 patients, which is only 38.6% of the original cohort. Thus, with this small cohort of patients, minor differences in terms of survival may be undetected.
Paper IV:

The frequency of ALK rearrangements in a Swedish NSCLC population is lower than previously reported. The two immunohistochemical assays evaluated demonstrated low sensitivity compared to the gold standard method fluorescence in situ hybridization.

Fluorescence in situ hybridization (FISH) is regarded as the gold standard in the detection of ALK rearrangements and has been the only method guiding treatment decisions in ALK positive NSCLC patients. Nonetheless, FISH is expensive and difficult to interpret. Therefore, ALK IHC is often used as screening tool in clinical practice. Recently, the FDA approved an IHC assay (Ventana ALK (D5F3) CDx Assay, Roche Diagnostics Limited, UK) as a companion test to identify ALK positive patients likely to respond to ALK inhibitor therapy [162]. In Paper IV, we therefore compared the detection of ALK-positive patients with the recently FDA-approved IHC assay together with an in house Dako IHC protocol, using the same antibody clone, as well as with FISH. This was performed in three large Swedish NSCLC cohorts (in total 851 patients) incorporating clinical parameters and gene expression data for 194 patients.

ALK status was first assessed on the TMAs for the three investigated cohorts. Tumors presenting ALK-positivity in at least one assay, or a contradicting result in one or two of the assays, were re-evaluated on whole tissue sections using all three methods (FISH, IHC Ventana and IHC Dako).

With FISH, ALK rearrangement was detected in 13/754 (1.7%) of the NSCLC samples. With the Ventana assay, 14/791 (1.8%) were considered ALK positive, and with the in house Dako protocol, 11/806 (1.4%) were classified as ALK positive. When comparing the two IHC assays, 15 cases were positive in at least one of the IHC assays but only ten (66.7%) cases were positive with both protocols. Comparison between FISH and the Ventana IHC protocol demonstrated nine positive cases with both methods, however, additionally four cases were rearranged according to FISH, but not positive with the Ventana protocol. Furthermore, five cases were positive with the Ventana protocol, but not rearranged according to the FISH assay. Comparison between FISH and the Dako IHC protocol demonstrated eight positive cases with both methods. Additionally, five cases were rearranged with the FISH assay, but these cases showed negative protein expression. Moreover, three additional cases displayed positive ALK protein expression when using the Dako protocol, but were FISH negative.
If FISH is considered the reference method, the sensitivity and specificity of the Ventana IHC assay was 69% and 99%, respectively, while for the Dako assay it was 62% and 100%.

To supplement the FISH and IHC assays, gene expression data from 194 patients (Uppsala I cohort) was available. Six cases were defined as gene expression positive, but only three of these were positive according to one of the three other ALK assays (FISH, Ventana IHC and Dako IHC). Notably, the two cases with the highest gene expression values were also positive with FISH.

When incorporating clinical parameters, the correlation analysis revealed known clinical association with lower age, female sex, and non-smoking for all assays, but the grade of association was significant only for smoking status. A tendency towards a longer overall survival for ALK positive patients was observed with the three assays when combining all histologies, but significance was not reached either in the uni– or in the multivariate analyses.

In conclusion, the results from this comparative study display a low frequency of ALK aberrations in the Swedish population of NSCLC patients. Even though a FDA approved IHC assay was used to detect ALK fusion proteins, a low sensitivity was obtained, therefore, the use of this IHC assay when screening for ALK-positive patients is questionable. The poor overlap between the different methods stresses the need for thorough validation of all ALK detection methods prior to implementation into clinical practice. Additionally, we demonstrated that there are patients with positive ALK protein expression and high mRNA levels that are FISH negative, perhaps specifying a distinct group of patients that also would benefit from ALK inhibitor therapy.
Concluding remarks and future perspectives

Genetic aberrations, leading to constitutive protein activation, have been identified in subgroups of lung cancer patients and some of these aberrations are today targets of specific inhibitors. Unfortunately, these inhibitors are effective only in small subsets of patients and the response is often only temporary. Therefore, further efforts are warranted to better understand the molecular background of NSCLC in order to find new treatment targets or to refine existing therapy.

The four papers presented in this thesis address this important aspect of clinical cancer research in different aspects – including evaluation of the expression of potential new treatment targets and prognostic biomarkers, and the evaluation of detection strategy of an already established treatment target.

In **Paper I**, ectopic expression of claudin 6 and claudin 18.2 was detected in small subsets of NSCLC patients. Since trials with monoclonal antibodies targeting claudin 6 (IMAB027) in ovarian cancer and claudin 18.2 (IMAB362) in gastro-esophageal cancer are ongoing (NCT02054351 and NCT01630083), these claudins may be targets in NSCLC as well. However, the effect of these antibodies has yet to be demonstrated. Furthermore, the frequency of claudin 6 and 18.2 expression in NSCLC is low, questioning a broader application of these therapeutic antibodies in this tumor entity. A demonstration of the functional importance of claudin 6 or 18.2 expression would be of interest for future studies, although antibody based therapy is not necessarily depending on the function of the protein.

In **Paper II** and **Paper III**, we described the expression of the pro-inflammatory mediator COX-2 and the effect of the specific inhibition of COX-2 on survival of NSCLC patients. Since tissue inflammation possesses pro-tumorigenic properties, anti-inflammatory strategies may present valuable alternatives to standard treatment. To date, the most frequently studied anti-inflammatory target is COX-2 [68]. Also in lung cancer, many studies have evaluated the impact of COX-2 expression on prognosis or investigated the effect of specific COX-2 inhibitors in the therapy of NSCLC patients [79,159-160]. With the results from Paper II and III, we propose that COX-2 expression is not associated with prognosis and is of minor importance regarding the treatment of NSCLC. It should be noted that other parameters
may be decisive for the response of COX-2 inhibitor therapy. For instance, the use of aspirin was associated with longer survival in patients with colorectal cancer harboring PIK3CA mutations [163]. Therefore, it would be interesting to analyze PIK3CA mutations in the tumor samples of the Cyclus cohort. Nonetheless, the minimal remaining tissue of these small biopsies makes this analysis difficult.

In Paper IV, we compared the results of different ALK-detection methods and analyzed the overall frequency of ALK rearrangements in NSCLC. Patients harboring ALK translocations display a remarkable response to ALK-inhibitor therapy, and it is therefore of importance to accurately identify patients benefitting from this therapy. Our results stress the difficulties associated with the reliable diagnostic identification of ALK rearrangement in NSCLC samples. This study lead to the change of our in house IHC protocol to the FDA approved Ventana protocol in the routine diagnostic. Furthermore, patients with clinical characteristics (non-smoker, younger age) indicative for ALK rearrangement should be evaluated directly using both FISH and IHC. Irrespective of the choice of analysis method we demonstrated a lower frequency of ALK rearrangements than previously reported. This stresses possible ethnic and geographical differences in the molecular background of NSCLC.

The challenge of detecting ALK rearrangements has led to the introduction of other techniques in the diagnostic setting. Polymer chain reaction (PCR)-based techniques can be used, but these assays rely on the knowledge of the fusion transcript variants, and will miss novel or rare translocations and would therefore not be optimal [164]. Other studies have suggested next generation sequencing methods [165-166] as well as exon array profiling in the detection of translocations [167]. One of the most promising techniques for future diagnostics may be the NanoString technology, which is based on direct multiplexed measurement of gene targets in a single reaction [168]. The discordant tissue samples of our present study present the optimal validation material for this new technique and analyses with the NanoString technology is planned in the near future.

As a valuable consequence of the efforts to characterize the described biomarkers in NSCLC, the previously established Uppsala cohort (Uppsala I) was further expanded with molecular data. Even more important, a complete new NSCLC cohort (Uppsala II) was established. Thus, both cohorts, including altogether 708 patients, are a unique source for further studies integrating genomic, gene expression and protein data.
Acknowledgements

This work was performed at the Department of Immunology, Genetics and Pathology (IGP) at Uppsala University, and the Molecular Pathology unit at the Department of Clinical Pathology, Uppsala University Hospital. Financial support was generously provided by the Lion’s Research Foundation, the Swedish Cancer Society and the Regional Lung Cancer Registry of the Uppsala-Örebro Region.

I would like to express my greatest gratitude to everyone who have encouraged and supported me during my PhD studies. I would especially like to thank the following persons:

All lung cancer patients for unselfishly donating their tissue for medical research.

My greatest gratitude to my supervisor Patrick Micke – thank you for everything! Thank you for sharing your wisdom, for believing in me, and for your patience and understanding. Thank you for putting things in perspective and for pushing me to do scary things. Thank you for giving me support and guidance when I needed it the most.

My co-supervisor, Johan Botling, thank you for your encouragement and your valuable comments throughout these years.

My second co-supervisor, Fredrik Pontén, thank you for always being so enthusiastic. Thank you for your valuable input in my projects.

Special thanks to the former member in our group, Karolina Edlund. Thank you for all your great ideas and valuable comments and feedback. Thank you for all brainstorming sessions and for putting up with all my questions.

I would like to thank the other members in our group. Firstly, Dijana, thank you for always supporting me, helping me put things in perspective and for your friendship. 😊 Secondly, Linnéa, thank you for being supportive and helpful throughout the years.
To all co-authors - thank you for good collaborations and for helping me with my research. It has been a pleasure working with you.

The Molecular Pathology group: Monica Lindell, Åsa Håkansson, Magnus Sundström, Lotte Moens and Ulf Thunberg, thank you for being so friendly and for making me feel like one in the group! Thank you for answering all my questions and for helping me in the lab.

Present and former room-mates at Rudbeck – thank you for providing such a friendly working environment, what would work be without you? 😊

Thank you Human Protein Atlas for letting me be a part of your research group. Thank you for your help with TMAs and scanning of slides!

Thank you IH-lab for all your help. I would especially like to thank Dijana Cerjan and Inga Hansson for always helping me and for putting up with all my immunohistochemistry-questions.

I would like to thank the whole Pathology Department here in Uppsala and especially Simin Tahmasebpoor for always being helpful and nice to me.

Thank you Christina Magnusson for all your encouragement! We PhD students can not wish for a better administrator!

Friends at Rudbeck! Thank you for your support and for sharing this fun but stressful time with me! 😊

Thank you friends outside of Rudbeck! Thank you for your encouragement and understanding! Now I hopefully will see you more often! 😊

Thank you Hanna! Without your support I don’t think I would have made it!

Thank you Simona for your great pep-talks! 😊 Thank you Monica for your support and for helping me with my tense shoulders! 😊 Thank you Jennifer for all your valuable comments and your great support! Thank you Mi for your enthusiasm, and for pushing me and helping me throughout these years! You all mean a lot to me!

Tack min kära familj. Tack Johan, mamma och pappa för att ni har stöttat och trott på mig alla år! Jag älskar er ♥

Daniel. Thank you for being so supportive and understanding. Thank you for always being there for me and for taking care of me during this stressful time. Jag älskar dig ♥
References


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-285844