Characterisation of *EGFR* and *KRAS* mutations in non-small cell lung cancer

Caroline Martinsson

*Supervisors: Karolina Edlund and Magnus Sundström*

*Department of Molecular Pathology*

*Uppsala University*
# Table of contents

Abstract ........................................................................................................................................... 3  
Sammanfattning .......................................................................................................................... 3  
Introduction .................................................................................................................................. 4  
Materials and methods ............................................................................................................... 8  
  Study population ....................................................................................................................... 8  
  DNA extraction .......................................................................................................................... 8  
  *EGFR* mutation analysis ......................................................................................................... 8  
  *KRAS* mutation analysis ........................................................................................................ 9  
Results ........................................................................................................................................... 10  
  Study population ....................................................................................................................... 10  
  *EGFR* mutation ..................................................................................................................... 10  
  *KRAS* mutation ..................................................................................................................... 11  
Discussion ..................................................................................................................................... 12  
  Conclusions .............................................................................................................................. 14  
Acknowledgements .................................................................................................................... 14  
References .................................................................................................................................... 15
Abstract

**Background:** Lung cancer is the leading cause of cancer-related death and one of the most common cancer types worldwide. Epidermal growth factor receptor (EGFR) has been shown to be an important therapeutic target in non-small cell lung cancer. Kirsten rat sarcoma viral oncogene homologue (KRAS) is a downstream signalling molecule in the EGFR pathway. Lung cancer patients with **EGFR** mutations respond to tyrosine EGFR inhibitor therapy, in contrast, patients with **KRAS** mutations do not benefit of such treatment.

**Methods:** This study investigates the frequency of **EGFR** and **KRAS** mutations in non-small cell lung cancer patients. Fifty-one lung cancer patients with primary non-small cell lung cancer diagnosed between 1995 and 2005 in the Uppsala-Örebro region were analysed by Sanger sequencing and Pyrosequencing to determine the mutation status of these genes.

**Results:** Five **EGFR** mutations were found in four patients (8%), two deletions in exon 19, one point mutation in exon 20 and two point mutations in exon 21. **KRAS** mutations were found in 12 patients (24%), ten codon 12 mutations and two codon 61 mutations.

**Conclusions:** This study confirms previous observations regarding the frequency of **EGFR** and **KRAS** mutations in non-small cell lung cancer. Mutations in **EGFR** and **KRAS** were mutually exclusive, indicating that both mutations present relevant tumorigenic genomic aberrations.

**Keywords:** NSCLC, mutation, TKI, Sanger sequencing and Pyrosequencing

---

**Sammanfattning**


DNA är den genetiska koden i vår arvsmassa. En gen är en sekvens av DNA som styr cellerna. Mutationer är förändringar som skett i DNAt. Mutationer kan antingen vara ärftliga och då befinner sig mutationen i könscellerna eller så är mutationerna inte ärftliga och befinner sig i andra celler (somatiska celler). Mutationer kan bidra till uppkomsten av cancer. I vissa fall är det viktigt att veta i fall en patient har en mutation i en viss gen.

Målinriktade terapier verkar på en specifik plats där en mutation har skett hos lungcancerpatienten. Exempel på mutationer som kan ske är till exempel i generna för **EGFR** och **KRAS**. För att ta reda på vilka patienter som kan ha nytta av målinriktade terapier utförs tester på DNA. Målinriktade terapier fungerar på de lungcancerpatienter med en **EGFR** mutation men är verkningslösa på de patienter med en **KRAS** mutation. Av 51 lungcancerpatienter upptäckades 8% med en **EGFR**-mutation och 24% med en **KRAS**-mutation. **EGFR** mutationer analyseras kliniskt i Uppsala medan **KRAS** mutationer hos lungcancerpatienter för tillfället inte analyseras i Uppsala.
Introduction

Lung cancer is one of the most common cancer types worldwide and is the leading cause of cancer-related death [1]. The 5-year survival rate is only 15% and has not increased significantly in recent years [2]. Approximately 3000 new patients are diagnosed with lung cancer each year in Sweden, among 700 from the Uppsala-Örebro region¹. About 54% of the 3000 newly diagnosed patients are men and 46% are women. The average age at time of diagnosis is 70 years. Lung cancer can be divided into two main categories, small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) [3]. In Sweden between 2002 and 2007, 79% of the lung cancer patients were diagnosed with NSCLC, 15% with SCLC and 6% were undefined¹. SCLC is the most aggressive form of lung cancer [3].

The introduction of targeted therapies has increased the survival time in a subset of patients with NSCLC. Mutations in the genes for epidermal growth factor receptor (EGFR) and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homologue (KRAS) may predict eligibility for EGFR inhibitor therapy and were among the genes previously analysed in the original patient cohort. The aim of present study was to analyse the frequency and distribution of KRAS and EGFR mutations in additional 51 NSCLC specimens that will be part of an extended study cohort.

NSCLC can be divided into three major subgroups based on histology: adenocarcinoma (50%), squamous cell carcinoma (30%) and large cell carcinoma (20%)¹. Figure 1 shows histological images of NSCLC patients. Adenocarcinomas arise from glandular cells in the lung mucosa and squamous cell carcinoma is located inside the bronchial tubes [3]. More men than women are diagnosed with lung cancer which closely follows smoking prevalence. Eighty percent of the lung cancer patients are diagnosed at an age above 60 [4]. Smoking is the leading cause of lung cancer; however 10% of lung cancer patients are non-smokers [5]. Other causes of lung cancer are environmental factors such as exposure to radon, asbest or air pollution [6]. The first symptomatic appearance of lung cancer is often quite vague and therefore patients often attend medical help too late. Symptoms that may indicate lung cancer include dyspnea (breathlessness), haemoptysis (coughing blood), chronic cough, chest or bone pain, weight loss, fever, depression, loss of appetite, and swallowing difficulties [3]. Many of these, symptoms like bone pain, fever and weight loss, are non-specific and particularly hard to diagnose in elderly people. Bronchoscopy and computed tomography are the methods of choice to diagnose lung cancer [3].

Figure 1. Histological images of fresh frozen tissues from NSCLC patients. (A) adenocarcinoma, (B) squamous cell carcinoma and (C) large cell carcinoma (C). All histology images are hematoxylin and eosin stained.

Regional Oncologic Center is responsible for regional cancer registration and contributes to the development of quality cancer care.
Tumours of the lung are classified according to tumour stage, on a scale ranging from I to IV, according to World Health Organisation (WHO) criteria and the TNM system. Tumour stage indicates the local extent and spread and represents until now the best marker for patient survival [7]. Patient with stage I–IIIa, i.e. localised stages, are eligible for surgical resection. Indeed only one-quarter of NSCLC patients undergo surgery aiming to remove the lung lobe or less often the entire lung [3]. However, even if the cancer is detected at an early stage and surgery is used as treatment, 30-55% of patients will relapse and die from metastatic recurrence within five years [9]. There is currently no way to identify those with risk of relapse. Patients with advanced disease (stage IIIb or distant metastases) are not operable and receive palliative treatment like radiation or chemotherapy. Chemotherapy can also be administered in combination with radiation therapy and may be given before or after surgery. Targeted therapies, that specifically inhibit the molecules or mechanisms that allow the tumour to grow or inhibit the blood supply of the tumour, are also available today [3]. Examples of targeted therapies are small molecule tyrosine kinase inhibitors (TKI) and monoclonal antibodies directed against the EGFR.

EGFR has been shown to be an important therapeutic target in NSCLC, as well as in colorectal cancer, head and neck squamous cell carcinoma and pancreatic cancer [10, 11]. EGFR belongs to the HER family of receptor tyrosine kinases [10]. This family includes EGFR/ERBB1, ERBB2, ERBB3 and ERBB4. The family members have an extracellular ligand binding region, a hydrophobic trans-membrane domain and a tyrosine kinase (TK) domain. Members of the EGF family of growth factor ligands can activate EGFR [12]. Ligand binding induces the formation of receptor homo- or hetero-dimers, followed by phosphorylation of specific tyrosine residues, resulting in activated intracellular signalling pathways [12, 11]. Examples of such signalling pathways are phospholipas C, phosphatidylinositol 3-kinase (PI3K)/AKT, mitogen activated protein kinase (MAPK) and signal transducer and activator of transcription (STAT) [10]. Aberrations in the EGFR signalling are considered to be associated with tumorigenesis in NSCLC.

Growth factor receptor tyrosine kinases can become oncogenic by a number of different events, including increased expression of the receptor ligand, activation of the receptor by mutation or over-expression of receptors. Therapies that focus on blocking growth factor receptor tyrosine kinases pathways have been developed, aiming to restrain tumour cell growth [11]. There are currently two different targeted therapeutic strategies available for EGFR inhibition, small molecule tyrosine kinase inhibitors and monoclonal antibodies. EGFR TKIs are small molecules that migrate through the cell membrane and attach to the ATP-binding site of the receptor. TKIs prevent activation of downstream targets and blocks the signalling pathways. TKI drugs that are currently used for NSCLC therapy are Gefitinib and Erlotinib [11]. Monoclonal antibodies directed against EGFR attach to the extracellular part of the receptor and prevent ligand binding. Examples of EGFR monoclonal antibodies currently in use for colorectal carcinoma are Cetuximab and Panitumab [11].

Only a minor group of NSCLC patients does clearly benefit from EGFR TKI therapy. Therefore it is important to identify markers to predict treatment outcome. Certain EGFR mutations have been found to correlate with clinical response to TKIs [11]. However less than 15% of NSCLC patients exhibit EGFR mutations [13]. Around 90% of these are deletions in exon 19 (e.g. E746-A750) or point mutations in exon 21(L858R). Less common (about 5%) are mutations in exon 18 and exon 20 [14-16]. Patients sometimes develop resistance to TKI therapy [17]. A resistant mutation (T790M) has been detected in exon 20. This mutation changes the structure of the TK domain and prevents efficient TKI binding [11].
RAS proteins are a large family of guanine tri-phosphate (GTP) binding proteins, encoded by the KRAS, neuroblastoma RAS viral oncogene homologue (NRAS) and Harvey rat sarcoma viral oncogene homologue (HRAS) genes [10]. These proteins are downstream signalling molecules in the EGFR pathway [18]. KRAS mutations were found in lung cancer as early as 1984 [19]. About 20% of NSCLC patients have a mutated KRAS, usually found in adenocarcinomas [10, 20]. Activating KRAS mutations are found in codon 12, 13 and 61: with 95% in codon 12 [21, 18]. These mutations result in a KRAS protein that is constantly activated independent from EGFR signalling. The tumour will therefore be resistant to therapy targeting the EGFR [11]. The relationship between activating KRAS mutations and lack of response to anti-EGFR monoclonal antibody therapy was first shown in colorectal cancer [10], where mutation analysis is today is routinely performed to guide EGFR antibody therapy (Cetuximab and Panitumab) [10].

In summary, three different lung cancer patient groups can be identified. The first group consists of patients with an EGFR mutation. These patients respond well to TKI therapy. The second group consists of patients with a KRAS mutation in which anti-EGFR therapies are not suitable. The third group consists of patients who have no mutations in neither EGFR nor KRAS, and where the response to therapy cannot be predicted today [10]. Therefore it is essential to accurately determine the patient mutation status before an EGFR-TKI can be applied.

Various techniques are currently utilized to detect mutations in clinical tumour samples. Most often Sanger sequencing is used, a method developed already in 1977 [22]. It is based on DNA synthesis with the addition of deoxynucleotide triphosphate (dNTP) and dideoxynucleotide triphosphate causing termination of newly synthesised DNA molecules [23]. First the DNA is amplified by conventional PCR and purified by removal of residual primers and nucleotides. Then the PCR products are again amplified with a specific sequencing primer and fluorescent terminating nucleotides that cause the reaction to stop resulting in fragments with different lengths. These fragments are then separated in a thin capillary and detected using a laser in which the various termination molecules emits a signal at different wavelengths. This leads to different peaks in four different colours, representing the four different nucleotides. Computer software determines the order of bases in the DNA region of interest. Sanger sequencing can analyse up to 10000 bases in a row [24].

Pyrosequencing represents another reliable alternative technique. Pyrosequencing is a method developed in Sweden, suitable for analysis of shorter DNA fragments than the Sanger sequencing, approximately up to 100 bases. It is a method based on real-time monitoring of DNA synthesis where each bound nucleotide will generate a light signal [23]. First the DNA is amplified by a PCR with one biotinylated primer. The biotinylated PCR products are then immobilised on streptavidin-coated Sepharose beads. A sequencing primer is mixed with the PCR product and binds to the template. DNA polymerase, ATP sulfurylase, luciferase, apyrase, adenosine 5’phosphosulfate (APS) and luciferin are added to the single-stranded template. Then the first dNTP is added. DNA polymerase binds to the DNA and if the added nucleotide is incorporated into the complementary chain, a pyrophosphate (PPi) is released. If the added nucleotide does not fit, no reaction will occur. When PPi is released sulfurylase phosphorylates PPi and APS to an ATP. ATP converts luciferase to an oxyluciferin and light is emitted [23]. The light is measured by a photometer, detected by a sensor and displayed as a peak in a graph (Pyrogram). The height of each peak is proportional to the number of incorporated nucleotides [24]. To remove any background, apyrase removes ATP and nucleotides that have not bound to the DNA template after each reaction cycle. New nucleotides will periodically be added. The complementary chain is elongated and the nucleotide sequence is determined by the peaks in the Pyrogram [23].
This project was a part of a collaborative study aiming to identify clinically relevant molecular profiles in NSCLC using a variety of methods, including array-based global gene expression analysis, SNP-array analysis for genome-wide detection of copy number aberrations, mutation analysis for selected cancer-related genes, fluorescence in-situ hybridisation (FISH) and analysis of protein expression using immunohistochemistry on tissue microarrays. Results from the molecular analysis are correlated to clinical patient parameters (survival, sex, smoking history, treatment response) and histopathological data (tumour type, differentiation), as well as to each other (gene expression to genomic alterations to mutation status etc). Until now, a total of 100 NSCLC samples have been included and are currently evaluated.
Materials and methods

Study population

Swedish lung cancer patients with primary NSCLC, diagnosed between 1995 and 2005, were identified in the regional lung cancer registry for the Uppsala-Örebro region. The register provides detailed information regarding patient diagnosis, histology, tumour stage, recommended treatment and progression status according to guidelines from the WHO. Patients with fresh frozen tissue available in the Uppsala biobank at the Department of Pathology and tissue specimen consisting of at least 50% tumour cells were included in the project. From all available tissue blocks in the Uppsala biobank at the Department of Pathology, hematoxylin-eosin (HE) stained cryosections were previously prepared and evaluated by a pathologist to ensure a tumour cell content of ≥50%. For this project 51 samples were analysed. Some tissue blocks were trimmed to obtain sufficient tumour cell content, and new HE-stained sections were made to confirm the quantity of tumour in the samples. All samples were used in accordance with the Swedish biobank legislation (Uppsala ethical review board reference # 2006/325).

DNA extraction

DNA was extracted from fresh frozen tissue samples using QIAamp DNA Mini Kit (Qiagen, Hamburg, Germany). Briefly, ten sections (10µm) were cut from each frozen tissue sample using a cryostate. Tissue sections were incubated in Buffer ATL (Qiagen) and proteinase K at 56°C until the tissue was completely lysed (1-3 hours). The DNA was then purified according to the manufacturer’s protocol (DNA Purification from Tissues; QIAamp DNA Mini Kit).

EGFR mutation analysis

Sanger sequencing was used to analyse EGFR exons 18, 19, 20, and 21. DNA was amplified by PCR (GeneAmp 2700 PCR system, Applied Biosystem) in 25µl reactions containing 15ng of template DNA, 22.5µl PCR MasterMix (Thermo Scientific) and 0.2µM of forward and reverse primers (table 1). PCR conditions were 40°C 10 min, 95°C 10 min, 35X (94°C 30s, 59°C 30s, 72°C 45s) and 72°C for 7 min. Amplification of a single PCR product of the expected size was electroforetically confirmed on a 2% agarose gel by ethidium bromide staining and UV-light. The PCR product was purified by removal of residual primers and nucleotides in reactions containing 6µl PCR product, shrimp alkaline phosphatase (1u) and exonuclease (5u) at 37°C for 30 min followed by 80°C for 20 min. The PCR product was then used as a template for sequencing reactions in both forward and reverse directions in 20µl reactions containing; 0.16µM forward and reverse primer, 1µl BigDye Terminator (Applied Biosystems) and 1µl purified PCR product. The sequencing conditions were 25X (96°C 10s, 50°C 5s, 60°C 4 min). The amplicon was then precipitated by a standard protocol using respectively 95% and 70% ethanol and 3M sodium acetate pH 4.6. DNA was diluted in 20µl formamide and analysed on an ABI 3130x1 Genetic Analyzer (Applied Biosystems).
Table 1. List of primer sequences in EGFR exon 18-21.

<table>
<thead>
<tr>
<th>Exon 18</th>
<th>Primer sequences in EGFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>GACCCTTGTCTCTGTGTCTTGT</td>
</tr>
<tr>
<td>Reverse</td>
<td>CCAGAGTCCCTTGAAGCTGTATA</td>
</tr>
<tr>
<td>Exon 19</td>
<td>Forward</td>
</tr>
<tr>
<td>Reverse</td>
<td>GCTGCTCTGCTCTAGACCCT</td>
</tr>
<tr>
<td>Exon 20</td>
<td>Forward</td>
</tr>
<tr>
<td>Reverse</td>
<td>AGATACGGGGAGGGGAGATA</td>
</tr>
<tr>
<td>Exon 21</td>
<td>Forward</td>
</tr>
<tr>
<td>Reverse</td>
<td>TTCCTGACACCAGGGGACCAG</td>
</tr>
</tbody>
</table>

**KRAS mutation analysis**

Pyrosequencing was used to analyse mutations in *KRAS* codon 12, 13 and 61, using the PyroMark Q24 *KRAS* Kit (Qiagen). In brief, DNA was amplified by PCR in 25µl reactions containing 10ng DNA, 1XPCR buffer (Invitrogen), 1.5mM MgCl₂, 0.2mM dNTP Mix, 1.25u Platinum Taq Polymerase (Invitrogen) and PyroMark Q24 *KRAS* forward and reverse primers for codon 12/13 or 61. PCR conditions were 95°C 15min, 40X (95°C 20s, 58°C 20s, 72°C 20s) followed by 72°C for 5 min. The biotinylated PCR products were immobilised on streptavidin-coated Sepharose beads. Sequencing primers for codon 12/13 and codon 61, respectively, were diluted to 0.1mM in Annealing Buffer (Qiagen) and 25µl of the diluted primer was used in each sequencing reaction on the PyroMark Q24 sequencer. All steps were performed strictly according to the instructions of the manufacturer.
Results

Study population

Lung cancer patients with primary NSCLC and available fresh frozen tissue in the Uppsala biobank at the Department of Pathology were included in the project. The tissue specimens contained at least 50% tumour cells. This study included samples from 51 NSCLC patients (male, 26; female, 25) with a mean age of 64 (range, 48-80) at time of diagnosis. Table 2 lists the baseline characteristics of the patients and analysis results for KRAS mutations and EGFR mutations.

Table 2. Baseline characteristics for all patients in the study and for patients with EGFR mutation and KRAS mutation.

<table>
<thead>
<tr>
<th></th>
<th>All patients</th>
<th>EGFR mutation</th>
<th>KRAS mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 60</td>
<td>17</td>
<td>33</td>
<td>2</td>
</tr>
<tr>
<td>≥ 60</td>
<td>34</td>
<td>67</td>
<td>2</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>Male</td>
<td>26</td>
<td>51</td>
<td>2</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>22</td>
<td>43</td>
<td>3</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>17</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>12</td>
<td>24</td>
<td>0</td>
</tr>
</tbody>
</table>

EGFR mutation

A total of five different EGFR mutations were found in four patients (8%, table 2). These included two deletions in exon 19 (E746-A750), one point mutation in exon 20 (V765M) and two point mutations in exon 21 (P848L and L858R, table 3). EGFR mutations were more common in adenocarcinomas than other histology tumour types: three were found in adenocarcinomas and one in squamous cell carcinomas. There were no differences between males and females (50% v 50%).
**KRAS mutation**

KRAS mutations in codon 12, 13 and 61 were analysed. Twelve KRAS mutations were found in 12 of 51 NSCLC patients (24%, table 2). Ten KRAS mutations were located in codon 12 (Gly12Val, n=3; Gly12Cys, n=3; Gly12Asp, n=3; Gly12Ala, n=1) and two mutations in codon 61 (Gln61His and Gln61Arg, table 3). KRAS mutations were more common among adenocarcinomas than in other histology tumour types: nine were in adenocarcinoma and three in large cell carcinomas. It was a difference regarding frequency of KRAS mutation between females and males (67% v 33%). KRAS mutations and EGFR mutations were mutually exclusive.

**Table 3.** Summary of final results of EGFR and KRAS mutations in non-small cell lung cancer patients.

<table>
<thead>
<tr>
<th>EGFR mutation</th>
<th>KRAS mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td><strong>KRAS codon 12</strong></td>
<td></td>
</tr>
<tr>
<td>Gly12Val</td>
<td>3</td>
</tr>
<tr>
<td>Gly12Cys</td>
<td>3</td>
</tr>
<tr>
<td>Gly12Asp</td>
<td>3</td>
</tr>
<tr>
<td>Gly12 Ala</td>
<td>1</td>
</tr>
<tr>
<td><strong>KRAS codon 61</strong></td>
<td></td>
</tr>
<tr>
<td>Gln61 His</td>
<td>1</td>
</tr>
<tr>
<td>Gln61 Arg</td>
<td>1</td>
</tr>
<tr>
<td><strong>EGFR exon 19</strong></td>
<td></td>
</tr>
<tr>
<td>Del E746-A750</td>
<td>2</td>
</tr>
<tr>
<td><strong>EGFR exon 20</strong></td>
<td></td>
</tr>
<tr>
<td>V765M</td>
<td>1*</td>
</tr>
<tr>
<td><strong>EGFR exon 21</strong></td>
<td></td>
</tr>
<tr>
<td>P848L</td>
<td>1*</td>
</tr>
<tr>
<td>L858R</td>
<td>1</td>
</tr>
</tbody>
</table>

* One of the four patients with EGFR mutation had two mutations in EGFR (V765M and P848L).
Discussion

Lung cancer is the leading cause of death from cancer worldwide [1]. EGFR has been found to be an important therapeutic target in NSCLC. EGFR can become oncogenic by activating mutations in its TK-domain [11]. Targeted therapies have been developed to block EGFR signalling pathways. There are currently two different targeted therapies, TKIs and monoclonal antibodies. TKIs prevent activation of downstream targets while monoclonal antibodies prevent ligand binding and/or dimerization to the EGF receptor [11]. TKI drugs currently in use for NSCLC are Gefitinib and Erlotinib: monoclonal antibody drugs currently in use for colorectal carcinoma are Cetuximab and Panitumab. Mutations in the EGFR and KRAS genes may predict suitability for EGFR TKI treatment. Not all NSCLC patients respond to EGFR TKIs and it is important to identify predictive markers for treatment benefit [11]. Patients with EGFR mutations respond significantly more often to EGFR TKI, while patients with KRAS mutations do not respond. Therefore the mutation status represents a valuable molecular marker for EGFR inhibitor therapy [11].

The aim of this project was to show the frequency and distribution of EGFR and KRAS mutations in 51 clinical NSCLC samples that will be part of an extended study cohort aiming to identify clinically relevant molecular profiles in NSCLC. The results from the smaller cohort will be compared with other previous studies regarding EGFR and KRAS mutations.

EGFR mutations were discovered in 8% of NSCLC patients. In the original study cohort consisting of 100 NSCLC samples, EGFR mutations were found in 13%. Rosell et al. identified EGFR mutations in 17% of 2105 highly selected lung cancer patients, mostly adenocarcinomas in women [13]. The slightly lower frequency of EGFR mutation in our cohort can be explained by the small samples size of the current cohort or by cohort composition with unselected patients of all histological subtypes.

Adenocarcinoma presents the most common histology (50%) \(^1\) of NSCLC with the highest frequency of EGFR mutations [20]. Rosell et al. showed that 81\% of the EGFR mutations were found in patients with adenocarcinomas and 70\% were found in females [13]. This study found the EGFR mutation in 75\% of patients with adenocarcinomas (table 2). There were no difference between males and females regarding EGFR mutation status. Rosell et al. also reported that more NSCLC patients with an age over 57 years had EGFR mutations [13]. In this study 67\% of the NSCLC patients at diagnosis had an age over 60 years. However no connection between age and EGFR mutation status was observed. In summary, NSCLC patients with adenocarcinoma have higher probability to have an EGFR mutation than those with other histologies.

In this study exon 18-21 of EGFR were analysed, although mutations in other exons exists. It is common to analyse only the hotspots in exon 19 (E746-A750) and 21 (L858R), as 90\% of all EGFR mutations are found in these exons [14]. Only nine percent of the EGFR mutations have been reported to be located in exon 18 and 20 [16].

Occasionally, more than one mutation is detected in the same NSCLC sample. In these cases, usually the two mutations are not equally significant in lung cancer. In this study two EGFR mutations in different exons were discovered in one sample, one point mutation in exon 20 (V765M) and one in exon 21 (P848L). Both these mutations are rare and their relevance in lung cancer is unknown. De Gunst et al. discovered that mutation P848L is likely

Regional Oncologic Center is responsible for regional cancer registration and contributes to the development of quality cancer care.
to be an uncommon and functionally silent EGFR polymorphism [25, 26]. Wu et al. showed that patients with a mutation V765M responded to Gefitinib [16].

It has previously been reported that KRAS and EGFR mutations are mutually exclusive in NSCLC patients [20]. In accordance with our study, EGFR mutations were not present in tumours with a KRAS mutation [20, 27]. The amount of KRAS mutations varies between 10-30% in different studies [28, 20]. In the previously analysed Uppsala cohort, KRAS mutations were found in 26% of the samples. This study found 24% KRAS mutations in 51 NSCLC patients, which confirms previously reported frequency of KRAS mutations in NSCLC.

The present study confirms that most of the KRAS mutations in NSCLC occurred in patients with adenocarcinoma [20]. This study found 75% of the KRAS mutations in patients with adenocarcinoma. Zhu et al. found 54% of the KRAS mutations in patients with adenocarcinoma. In their study 57% of the patients with KRAS mutation were over 60 years of age at diagnosis but no significant differences between males and females was seen [21]. In our cohort 67% of the patients with a KRAS mutation were females and 67% were above the age of 60 at diagnosis. In summary, females with an age over 60 years and with adenocarcinoma have a higher probability to develop KRAS mutations.

The sensitivity of the methods is an important factor in molecular diagnostics. Human tissues may contain a low percentage of tumour cells. When a sample contains too few tumour cells in a background of normal cells, this will reduce the chance to detect a mutation. In this study we only selected samples with at least 50% tumour cells. The Uppsala molecular pathology laboratory has three methods for the analysis of EGFR and KRAS mutations, Sanger sequencing, Pyrosequencing and real-time PCR based DxS TheraScreen method. We choose to analyse KRAS mutations with Pyrosequencing, as it is a well-established and rapid method. Pyrosequencing has sensitivity to detect mutations even if it exists only in 2.5% of the cells in a sample [29]. Pyrosequencing is suitable for KRAS mutation hotspots in codons 12, 13 and 61. So far, there is no established protocol to identify EGFR mutations with Pyrosequencing. The difficulties lie in the spreading of the EGFR mutations among various exons. Further, Pyrosequencing is not optimal to identify deletions. Therefore we choose to determine EGFR mutations using Sanger sequencing which have a sensitivity to detect mutations if they exists in 25-30% or more of the cells in a sample [30]. The method is more time consuming than Pyrosequencing, but the results represent a comprehensive overview over the exons. DxS Thera Screen is a very sensitive method that detects a mutation even if it only exists in 1% of the cells in a sample according to the TheraScreen manual. The method is based on real time PCR and detects specific mutations in KRAS and EGFR. In EGFR, DxS detect specific hotspots in exon 18 (G719A/S/C), deletions in exon 19, point mutation (T790M) and insertions in exon 20 and point mutations (L858R, L861Q) in exon 21. In KRAS, DxS detect specific hotspots in codon 12 (Ala, Asp, Arg, Cys, Ser and Val) and only one in codon 13 (Asp). DxS is the method of choice to find predefined EGFR and KRAS mutations, but is expensive. In our study there was no need for such a sensitive method, as we only used samples with at least 50% tumour cells. The Uppsala molecular pathology laboratory applies this method only when samples with low tumour content are analysed. Since the real frequency and the clinical impact of mutations are important, other analysis than the known hotspots in exon 18 to 21 is preferable. For instance DxS do not detect any mutations in codon 61 in KRAS. In this study two of 12 mutations in KRAS were located in codon 61 (table 3). In summary, as we only used samples with at least 50% tumour cells our choice to apply Pyrosequencing to identify KRAS mutations and Sanger sequencing to detect EGFR mutations was reasonable.
Different studies discuss the best way to find responders to EGFR TKI, *EGFR* mutations or increased *EGFR* gene copy number. FISH is a method using probes with fluorescent labels, in order to evaluate the intact cells for genetic changes [31]. *EGFR* FISH is a method for *EGFR* gene copy number evaluation. Cappuzzo et al. compared *EGFR* mutation and *EGFR* FISH from different studies [32]. His conclusion was that both *EGFR* FISH and *EGFR* mutations are good markers for the response to EGFR TKI therapy, although Cappuzzo concluded that *EGFR* mutations are a better marker for the identification of the responders. *EGFR* FISH has been proven to be a better assay to measure a patient’s survival rate when exposed to TKI [32].

Recently a novel fusion protein was identified in a small subgroup of NSCLC, a fusion between the echinoderm microtubule-associated protein-like 4 (EML4) with the intracellular kinase domain of the anaplastic lymphoma kinase (ALK). EML4-ALK translocation and *EGFR* and *KRAS* mutations are mutually exclusive. Patients with EML4-ALK are considered to be resistant to EGFR TKI therapy [1]. Instead, astonishing response to a novel ALK inhibitor was observed in a clinical phase II in EML4-ALK fusion-protein positive NSCLC patients [33].

**Conclusions**

This study found four (8%) NSCLC patients with *EGFR* mutations that, had a high probability to benefit from EGFR TKI therapy. No NSCLC patient with *EGFR* mutation revealed a *KRAS* mutation. The results from this study will be incorporated in a larger study.

Next step in this study is to analyse ELMK4-ALK fusion protein by FISH, immunhistochemical analysis of the patient cohort on tissue micro-array, and to include SNP-array analysis. The markers from both molecular studies will be correlated to clinical patient and histopathological parameters. The results will comprehensively characterise a well defined patient cohort with reliable molecular aberrations and may increase the general knowledge on lung cancer tumorigenesis.

**Acknowledgements**

I would like to thank my supervisors Karolina Edlund and Magnus Sundström for all the help I received throughout the project. In addition, I would like to thank Patrick Micke, Johan Botling and everyone who works at the laboratory of clinical molecular pathology, for showing interest and wanting to help me throughout the project.
References


