THERAPEUTIC TARGETS AND BIOMARKERS IN LUNG CANCER

Mariëlle I Gallegos Ruiz
The research described in this thesis was performed at the Department of Medical Oncology, VU University Medical Center, Amsterdam The Netherlands. The study was supported by the KWF Dutch Cancer Society by Grant VU-2005-3445 and by Roche Nederland B.V. Additional support was obtained from mr. Anne de Haan.

Nederlandse Titel:
Therapeutische targets en biomarkers in long kanker

Publication of this thesis was financially supported by:
KWF Dutch Cancer Society, Roche Nederland B.V. and BD Biosciences.

Cover: Dirk van de Ven
Printed by Ipskamp Printpartners, Enschede The Netherlands

© Copyright Mariëlle I. Gallegos Ruiz, Amsterdam, 2008
All rights reserved. No parts of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronically, mechanically, by photocopying, recording or otherwise without the written permission of the author.
ISBN: 978 90 8659 193 0
VRIJE UNIVERSITEIT

THERAPEUTIC TARGETS AND BIOMARKERS IN LUNG CANCER

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
on gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op dinsdag 22 april 2008 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Mariëlle Ilse Gallegos Ruiz
geboren te Utrecht
promotoren:  prof.dr. G. Giaccone
            prof.dr. E.F. Smit

copromotoren:  dr. J.A. Rodríguez
                   dr. B. Ylstra
CONTENTS

CHAPTER 1  General Introduction  7
CHAPTER 2  Erlotinib for frontline treatment of advanced non-small cell lung cancer: a phase II study  29
CHAPTER 3  Response to Erlotinib in first-Line treatment of non-small cell lung cancer in a white male smoker with squamous-cell histology  47
CHAPTER 4  Genetic heterogeneity in patients with multiple neoplastic lung lesions: a report of three cases  57
CHAPTER 5  EGFR and K-ras mutation analysis in non-small cell lung cancer: comparison of paraffin embedded versus frozen specimens  77
CHAPTER 6  EGFR gene copy number detection in non-small cell lung cancer; a comparison of fluorescent in situ hybridization and chromogenic in situ hybridization  91
CHAPTER 7  Functional analysis of cancer-associated EGFR mutants using a cellular assay with YFP-tagged EGFR intracellular domain  105
CHAPTER 8  Combined assessment of EGFR pathway-related molecular markers and prognosis of NSCLC patients  123
CHAPTER 9  Integration of gene dosage and gene expression in non-small cell lung cancer; identification of HSP90 as potential target  145
CHAPTER 10  Summarizing discussion and future perspectives  163
APPENDIX A  Supplementary information Chapter 8  177
APPENDIX B  Supplementary information Chapter 9  185
APPENDIX C  Supplementary information Colored figures  195

NEDERLANDSE SAMENVATTING  211
CURRICULUM VITAE  219
DANKWOORD  225
CHAPTER 1

GENERAL INTRODUCTION
LUNG CANCER

Lung cancer is the leading cause of cancer-related deaths worldwide [1], there are approximately 1.2 million new cases of lung cancer diagnosed every year worldwide, and 1.1 million patients die of the disease. The main cause of lung cancer is exposure to tobacco smoke [2]. In Europe and US lung cancer in never-smokers is rare (<10% of cases), and most cases in the US are now represented by ex-smokers. In non-smokers lung cancer causes can be radon gas, asbestos or genetic factors [3-5]. The most common symptoms of lung cancer at onset are malaise, weight loss and dyspnea. The histological or cytological diagnosis of lung cancer is usually obtained by bronchoscopic examination or fine needle aspiration.

Two main distinctions are made among the several histological types of lung cancer: small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC). Staging of SCLC is a simplified two class system, that classifies the extent of the disease into limited or extensive stage [6]. For NSCLC the Tumor-Node-Metastasis (TNM) classification is used for staging purposes [7]. This system was initially introduced in 1972 and has been modified over the years. Recently, some additional modifications have been proposed [8]. The TNM staging system includes a pretreatment clinical classification (cTNM) and a postsurgical histopathologic classification (pTNM). The cTNM is used for the guidance of treatment and the pTNM is used to estimate prognosis and to select the most appropriate treatment approach after surgery. The TNM staging system takes into account the size and the degree of spread of the primary tumor (T), whether and to what extent lymph nodes are involved (N) and the presence of distant metastasis (M). NSCLC accounts for 85% of all lung cancers and represents several different histological subtypes, the most important of which are squamous cell lung carcinoma, adenocarcinoma and large cell carcinoma. Whereas squamous cell carcinoma was the most frequent histological type a decade ago, adenocarcinoma represents now over 50% of NSCLC in the US and several North European countries and Japan. This change in histological prevalence is most likely related to a change in type of cigarette smoke and potentially to pollution in urban areas [9].

Prognosis for both SCLC and NSCLC is poor. The overall five-year survival for patients with limited stage SCLC is about 20% and for patients with extensive stage SCLC the five-year survival is less than 1% [6]. For NSCLC, the five-year overall survival of patients with stage IA is 67% and for patients with stage IIA this is 55%. For patients with stage IIIA disease, 23% is still alive five years after surgery, whereas for patients with stage IV this is only 1% [10].
CHAPTER 1

TREATMENT FOR LUNG CANCER

Treatment for lung cancer depends on the histological type (SCLC, NSCLC), the stage of disease, and performance status. For SCLC, chemotherapy represents the mainstay treatment since this type of lung cancer is usually metastatic at presentation and, therefore, surgery is rarely applied. The most common chemotherapy used in SCLC is a combination of cisplatin and etoposide [11]. Response to chemotherapy is as high as 90% in patients with limited disease, and 50-60% in patients with extensive disease. 10-50% of responders achieve a complete radiological response. In patients with limited disease, chest radiotherapy is given concomitantly to chemotherapy. Prophylactic cranial irradiation is recommended in patients who achieve a response to systemic treatment, in order to prevent relapse in the brain. The brain in fact represents a common and serious relapse site in patients with SCLC.

For early stage NSCLC, surgery is the most important curative modality. However, only approximately 20-30% of patients are diagnosed at resectable stage (stage I-II). For advanced NSCLC patients in good performance status (0-1), platinum-based chemotherapy is the standard treatment and partial responses can be achieved in 30-40% [12]. Complete responses are very rare in advanced NSCLC. Regimens used are cisplatin or carboplatin, in combination with gemcitabine, paclitaxel, docetaxel, etoposide or vinorelbine. Chemotherapy and radiotherapy can also be used as neo-adjuvant therapy with the intent to shrink the tumor before surgery or as adjuvant therapy to improve outcome after surgery. The use of both neo-adjuvant [13,14] and adjuvant chemotherapy [15,16] has been shown to improve patient survival. Despite advances in these combined treatment modalities for lung cancer, prognosis remains poor and severe side effects are often observed. Therefore, more effective and less toxic treatments are needed. For the treatment of advanced NSCLC, a variety of molecular targeted therapies have been recently introduced.

TARGETED THERAPY

In order to kill tumor cells and spare normal cells, targeted therapy aims at specifically targeting molecular structures or abnormalities present on tumor cells and absent or less abundantly present on normal cells. Three targeted agents have been introduced recently in the treatment of advanced NSCLC: Gefitinib (Iressa) and Erlotinib (Tarceva), that target the tyrosine kinase domain of the epidermal growth factor receptor (which will be discussed in detail in the next paragraphs), and Bevacizumab (Avastin) which targets the vascular endothelial growth factor [17]. Several other targeted agents are actively being investigated in lung cancer, such as the proteasome inhibitor Bortezomib [18], the BRAF inhibitor Sorafenib [19], the mTOR inhibitor rapamycin [20] and many others [21].
We will further focus on the EGFR TK inhibitors and will first explain the EGFR pathway in more detail.

**EGFR PATHWAY**

The epidermal growth factor receptor (EGFR) is a member of the ErbB family of receptors, a subfamily of four closely related receptor tyrosine kinases: EGFR (HER1; ErbB-1), HER2/c-neu (ErbB-2), Her 3 (ErbB-3) and Her 4 (ErbB-4). EGFR is overexpressed in many solid tumors including NSCLC [22]. The EGFR is present on the cell surface and is activated by binding its specific ligands, including epidermal growth factor and transforming growth factor alpha. Upon activation, EGFR undergoes a transition from an inactive monomer to an active dimeric form. In addition to forming EGFR homodimers, EGFR can pair with other members of its family to form activated heterodimers. Upon EGFR dimerization, the intracellular tyrosine kinase is activated, resulting in autophosphorylation of several tyrosine residues. This autophosphorylation subsequently results in downstream activation and signaling by several other proteins that associate with the phosphorylated tyrosines through their phosphotyrosine-binding SH2 domains. These downstream signaling proteins initiate several signal transduction cascades, mainly the PI3K/AKT, STAT and RAS/MAPK pathways, which regulate biological responses such as proliferation, cell motility, angiogenesis, cell survival and differentiation (Figure 1) [23,24]. Aberrant signaling by EGFR can lead to uncontrolled cell growth resulting in cancer. For this reason, EGFR inhibition is intensively being explored as a cancer therapeutic approach.
EGFR INHIBITION

Various approaches to inhibit EGFR are being investigated. Monoclonal antibodies block the extracellular domain of EGFR, thereby competing with binding of the ligand, resulting in the inhibition of activation of the receptor. An example is Cetuximab, which is approved for the treatment of advanced colorectal cancer [25]. Another approach to inhibit EGFR is using small molecules to inhibit the EGFR tyrosine kinase activity, thereby inhibiting the activation of signaling cascade on the cytoplasmic side of the receptor. Gefitinib and Erlotinib are two examples of tyrosine kinase inhibitors approved in several countries for the treatment of advanced NSCLC [26].

In NSCLC, responses to EGFR targeted agents are generally low in unselected patients, with response rates usually below 10% in Caucasians. The monoclonal antibody Cetuximab has shown modest activity in lung cancer [27]. A large randomized study, which compared cisplatin-vinorelbine vs cisplatin-vinorelbine-Cetuximab, revealed that the primary endpoint of improving overall survival was achieved with the addition of Cetuximab, indicating that probably the effect of combining this antibody with chemotherapy is at least additive [28].

Response rates to the EGFR TK inhibitors Erlotinib and Gefitinib are also low in unselected patients; initial studies showed responses in 10-20% of patients when used as second- or third line treatment for advanced disease [29-31]. Higher responses to EGFR tyrosine kinase inhibitors (TKIs) have been seen in east-Asian patients, in women, in non-smokers, and in patients with adenocarcinoma [32,33]. Erlotinib administered to patients after first or second line treatment with chemotherapy showed a 2 months prolonged survival compared to patients treated with placebo in the randomized BR21 study [34]. On the other hand, Gefitinib failed to improve survival compared to placebo treated patients in the ISEL study [35]. Explanations for these findings have been suggested to be caused by the difference in design of the study or the different dosing schedules. Whereas Erlotinib was administered at the maximum tolerated dose (150mg/d), Gefitinib was administered below its maximum tolerated dose (250mg/d). In four randomized trials comparing the addition of Gefitinib (INTACT1 and 2) or Erlotinib (TALENT and TRIBUTE) to platinum based doublet chemotherapy (gemcitabin/cisplatin or carboplatin/ paclitaxel) no benefit in survival was observed in patients receiving chemotherapy with either Gefitinib or Erlotinib compared to chemotherapy alone [36-39].

EGFR TK MUTATIONS AS PREDICTOR FOR RESPONSE TO EGFR TKIs

Among the patients showing a marked response to the agents Gefitinib or Erlotinib, it has been shown that approximately 80% of patients carry specific mutations in the tyrosine kinase domain
of EGFR [40-42]. These specific EGFR mutations are present in tumors of 10% of Caucasian patients and 30% of NSCLC patients with east-Asian ethnicity [43]. In the afore mentioned randomized trials (IDEAL, INTACT I-II and TRIBUTE) the effect of EGFR mutations has been determined retrospectively. EGFR mutations were correlated with the previously identified clinical characteristic related to EGFR TKI sensitivity (never smoking status, female gender, Asian ethnicity and adenocarcinoma histology). In the IDEAL trial, 46% of patients harboring an EGFR mutation showed a response to Gefitinib [44]. The INTACT trial analysis failed to show a correlation between Gefitinib response and EGFR mutations. Overall, Gefitinib did not show to improve survival in patients harboring EGFR mutations when given in combination with chemotherapy [44]. Molecular analysis of the TRIBUTE study showed an improved response rate to Erlotinib in patients with EGFR mutations, 53% of patients harboring an EGFR mutation showed a response to Erlotinib in combination with chemotherapy [45]. The absence of a response to Erlotinib in EGFR mutant patients could have possibly been the result of antagonism between chemotherapy and Erlotinib. For example, it has been reported that exposure of cells to pemetrexed followed by Erlotinib has a synergistic effect, although Erlotinib followed by pemetrexed can have an antagonistic effect [46]. An ongoing discussion is whether EGFR sensitizing mutations are predictive of response to EGFR TKI only or they represent positive prognostic markers in general. In a recent study with over 1600 patients randomized to Gefitinib or docetaxel in second or third line treatment, patients showing the clinical characteristics typically related to good response to EGFR TKIs, and EGFR mutations and copy number showed better survival in both arms of the study. The results of this study [47] are provocative and will need to be confirmed by other studies. Nearly 90% of the EGFR mutations detected are short deletions in exon 19 affecting the amino acids ELREA (Del746-750), or mis-sense mutations in exon 21 replacing leucine by arginine resulting in the amino acid change L858R. EGFR TK mutations have been studied in much detail by in vitro analyses and they have been shown to specifically activate anti-apoptotic pathways by activation of AKT and STAT pathways [48,49]. In vivo assays furthermore show that persistent EGFR mutant signaling is required for tumor maintenance [50,51]. These reports show that expression of mutant EGFR led to the development of adenocarcinomas in mice and that by reducing the expression of the transgene or by treatment with Erlotinib, the tumors regressed. It has been stressed that various EGFR mutations may have differences in sensitivity to EGFR TKIs. The EGFR deletion mutation in exon 19 has been shown to be superior to other EGFR mutations in predicting the response to TKIs [52]. It has also been shown that patients harboring this type of EGFR mutation have a significantly longer survival time than patients harboring the L858R mutation after treatment with Erlotinib or Gefitinib [53]. Moreover, the G719S mutation was much less sensitive to Gefitinib than the L858R mutation [54]. An insertion mutation in EGFR
exon 20 furthermore showed association with resistance to the EGFR TKIs [55]. To date, over 75 different mutations in the tyrosine kinase domain of EGFR have been reported. The hotspots for EGFR mutations in the tyrosine kinase domain are indicated in Figure 2 [56].

In addition to sensitivity mutations, a secondary resistance mutation has been identified in EGFR, replacing methionine for threonine at residue 790 (T790M). This mutation is observed in ~50% of tumors from patients showing acquired resistance to EGFR TKIs [57]. The T790M mutation is also rarely observed in untreated tumor samples. This indicates that there is a possibility that resistant subclones containing this additional mutation may exist, which get the chance to expand after the bulk of the tumor regresses as a consequence of the drug sensitivity conferred by the primary sensitizing mutation. A family with the lung carcinomas, of the broncho-alveolar carcinoma subtype, has also been described to contain a germline T790M mutation [58]. To overcome resistance by this T790M mutation, second generation TKIs are currently being developed and tested in the clinic [59,60].

Methods of assessing EGFR mutations have differed in the various studies published to date. The most commonly used method is a combination of PCR and direct sequencing [61]. It has, however, been claimed that single stranded conformation polymorphism gel electrophoresis is more sensitive than direct sequencing. This technique can, moreover, be used to screen a large number of samples [62]. Other methods under study are, among others, mutation enriched sequencing [63], real time Taqman PCR [64], and DNA endonuclease surveyor [65]. Highly sensitive methods might be able to detect subclones of cells harboring an EGFR mutation, but the significance of these low abundant mutants is uncertain.

Although EGFR mutations have been shown to predict response to TKI treatment in a subset of NSCLC patients, the correspondence is not perfect [66-68]. Prospective randomized studies that
select tumors of patients based on clinical characteristics and/or molecular markers are presently underway [69], and these studies will help to clarify several of the open questions.

ADDITIONAL PREDICTIVE PARAMETERS FOR RESPONSE TO EGFR TKIs

The correlation between the presence of classical EGFR mutations and sensitivity to EGFR TKIs is not 100%, and it has been shown that patients initially responding to EGFR TKIs can develop acquired resistance, which in ~50% of cases can be explained by a second T790M mutation. These issues highlight the need to identify additional markers related to sensitivity and resistance to EGFR TKIs.

One such marker is EGFR copy number. Various groups have reported that EGFR copy number is a better predictive marker for response or “benefit” of NSCLC patients treated with EGFR TKIs than EGFR mutations. It was shown that high polysomy and amplification of the EGFR gene predicted a favorable response to Gefitinib and high EGFR DNA copy number was also related with prolonged survival [70,71]. In the BR21 study, high EGFR DNA copy number was associated with longer survival in patients treated with Erlotinib [72]. Moreover, molecular analysis of the IDEAL/INTACT trials showed that patients with high EGFR copy number showed higher response rates than patients having a normal EGFR copy number (29% vs. 9%, respectively). EGFR amplification was not significantly correlated with female gender, never smoking status, adenocarcinoma subtype and Asian ethnicity [44]. This indicates that EGFR amplification can identify an additional subset of patients likely to respond to EGFR TKIs. Another study failed to show a correlation of EGFR copy number and response to EGFR TKIs, although it must be stressed that in this study a different technique (RT-PCR rather than FISH) was used to determine EGFR copy number [73]. In summary, these data indicate that EGFR gene copy number as assessed by FISH can be an important predictive biomarker for NSCLC patients treated with EGFR TKIs.

Another marker that has been linked to responsiveness to EGFR TKI sensitivity is phosphorylation of the downstream effector AKT [74]. pAKT was reported to be predictive for response to Gefitinib, and was also shown to be correlated with never smoking status, female gender and Asian ethnicity. pERK was also investigated but was not correlated with any of these findings. In another study, neither pAKT nor pERK were associate to response to Gefitinib [66].

KRAS mutations are detected in 20-30% of NSCLC patients, mainly in adenocarcinomas, and are generally related to poor prognosis [75]. In patients treated with EGFR TKIs, mutations in KRAS have been shown to be related to primary resistance to these agents [76,77]. KRAS mutations and EGFR mutations are usually mutually exclusive, and the presence of a KRAS mutation is a useful marker for selecting patients who will not benefit from EGFR TKI therapy.
MET encodes a receptor tyrosine kinase, capable of activating similar pathways (PI3K, MAPK and STAT) as are activated by EGFR. MET overexpression has been implicated in the development and progression of various human cancers [78]. In NSCLC, expression of MET has been shown to be a poor prognostic factor [79]. Amplification of the MET receptor has recently been linked to resistance to Gefitinib by activating the EGFR family member, Her3 [80]. Another recent study also identified MET amplification as an important marker related to acquired resistance to Gefitinib [81]. This latter study used array comparative genomic hybridization to identify differences in copy number changes between patients who developed acquired resistance to Gefitinib compared to EGFR mutated patients never treated with Gefitinib. They identified MET to be amplified in 21% of tumors who developed acquired resistance to Gefitinib, whereas MET was amplified in only 3% of patients never treated with Gefitinib. A subset of these MET amplified patients having acquired resistance to Gefitinib contained the EGFR mutation T790M. MET amplification thus identifies a new subset of patients with acquired resistance, and inhibition of MET could be an alternative strategy to treat these patients.

RNA AND DNA MICROARRAY TECHNOLOGY

A combined analysis of EGFR and its downstream molecules could likely identify additional markers for prediction or prognostic purposes. Inhibition of EGFR and its downstream molecules is an effective strategy to improve clinical management of a subset of NSCLC patients. However, it is likely that in addition to the extensively studied EGFR-related molecules, there are more genes important in NSCLC, that have yet to be identified.

Microarray analysis represents a powerful approach to identify pivotal genes, not previously recognized, in various diseases. By microarray technology the whole genome can be screened for aberrations at the RNA or DNA level. A microarray slide contains a large collection of DNA sequences that represent single genomic elements. By hybridizing fluorescently labeled RNA or DNA to these slides, one can measure the hybridization intensity and, thereby identify gene expression and copy number changes. The most commonly used commercial platforms nowadays are Affymetrix, Illumina and Agilent. Microarrays can be used for expression profiling, in which the expression levels of thousands of genes can be measured concurrently (RNA level). Microarrays can also be used for comparative genomic hybridization, array CGH. In array CGH, one is able to determine alterations in copy number changes along the entire genome (DNA level) [82].

As mentioned, microarrays can be used to identify potentially important molecules, for example by screening for individual genes or copy number changes differentially apparent between cancer patients with long and short survival. Microarrays can also be used to learn about complexity in
cancer biology by studying the gene expression or copy number changes under various experimental conditions. Yet another application of microarrays is to classify patients according to their gene expression or copy number profiles, thereby being able to identify clinically relevant subgroups of patients.

**FIGURE 3. Principle microarray technology.** Test (tumor) RNA or DNA and reference RNA or DNA are fluorescently labeled with i.e. Cy3 and Cy5 respectively. The test and reference samples are simultaneously hybridized to microarray slides containing tens of thousands of DNA oligonucleotides. The fluorescence intensity ratio between Cy3 and Cy5 can be measured and translated to gene expression or copy number changes of single genes in the whole genome.

**GENE EXPRESSION PROFILING IN NSCLC**

Using gene expression profiling, the expression levels for thousands of genes is explored simultaneously, which can be used, for example, to identify disease genes by comparing gene expression in cancer patients with and without recurrence.

Generally, there are two ways of analyzing gene expression data; unsupervised and supervised analysis. In unsupervised analysis, the gene expression ratios are used to define (clinically relevant) patient groups. On the other hand, a supervised analysis method can be used, in which
two groups with known outcome, for example recurrence, are analyzed to what extent they vary in gene expression profiles. In this latter way, gene signatures can be developed or gene targets can be identified.

Several molecular classifications of lung cancer based on global gene expression profiling have been published [83-85]. Many RNA gene signatures have been correlated to survival or recurrence [86-92]. Some of these signatures have been proposed to be of guidance in the use of adjuvant therapy. A rather surprising finding is the relative lack of overlapping genes present in the several signatures that have been developed. However, it does appear that the genes of the various signatures generally belong to the same pathways. A signature identified to predict outcome of lung adenocarcinomas has also been shown to be able to predict outcome in breast adenocarcinoma [93]. Indicating some signatures could also identify pathways responsible in common types of cancer.

One of the pitfalls of gene expression profiling studies is that a large amount of data easily allows the recognition of a certain pattern, without necessarily meaning this signature is of clinical importance. Therefore, multiple testing correction and validation of results is an important issue in analyzing gene expression microarray data.

COMPARATIVE GENOMIC HYBRIDIZATION IN NSCLC

Array comparative genomic hybridization (array CGH) is used to assess genome wide copy number changes [94,95]. Amplification and deletion of genes at the DNA level frequently contribute to tumorigenesis of NSCLC, and array CGH can be used to better define the complexity of this disease.

In analyzing array CGH data, normalization and approaches to reduce measurement of noise are of importance. Various algorithms have been developed and use the positional information from neighboring clones to identify breakpoints and segment the data [96,97]. Another important aspect in analyzing array CGH data is the calling of gains and losses, an example of an algorithm covering this feature is CGH call [98].

In NSCLC, several genomic regions have been described to frequently harbor DNA gains (3q, 5p, 7q, 8q, 11q and 16p) or losses (3p, 4q, 5q, 6q, 8p 9p and 13q, 17q) [99]. Characterization of these DNA copy number changes is important for both the basic understanding of NSCLC and also to identify recurrent patterns for classification or identification of targets for therapy.

Based on array CGH data in NSCLC, it has been shown that multiple molecular carcinogenesis pathways exist that are most likely related to gender and smoking habits [100]. Furthermore, array CGH can also give a refined classification of lung cancer based on genetic aberrations rather than on morphologic subtypes [101]. It has been shown that there is a large overlap
between aberrations observed in the adenocarcinoma and squamous cell carcinoma subtypes, except for 3q gains which seem to be more specific for the squamous cell carcinoma subtype [102].

By using array CGH for target discovery, one approach is the integration of array CGH and gene expression data to identify potential candidate genes [103,104]. In this respect, two focal amplifications of 8p12 and 20q11 have been studied in detail leading to the identification of two candidate genes (WHSC1L1 and TPX2), important in both NSCLC and ductal pancreatic cancer [102]. Using this approach, however, genes that may be over- or underexpressed because of other reasons, such as micro RNA regulation, methylation or mutation, remain unidentified [105].

OUTLINE OF THIS THESIS

The prognosis of non-small cell lung cancer remains poor, mainly due to advanced stage of the disease at the time of diagnosis. Systemic chemotherapy increases survival, but its effect is modest. The introduction of targeted agents in this disease has started improving the outcome of selected patients. For example, a specific subset of patients having never smoked, being of East Asians race, female and having adenocarcinoma histology, have a marked better outcome when treated with EGFR TKIs. Specific mutations in EGFR can also partially predict EGFR sensitivity to these agents. The aim of this thesis was to investigate the EGFR pathway, including EGFR mutations, in patients with Non-Small Cell Lung Cancer. In addition to the EGFR pathway, this thesis aimed to identify novel markers for prediction of response or prognosis purposes, and also to identify genes as candidate therapeutic targets for NSCLC.

In CHAPTER 1, a general introduction about lung cancer and current treatment approaches is given. This chapter focuses on the EGFR pathway and highlights some additional candidates for prediction biomarkers. The microarray technique is introduced to show how this technique can be used to search for additional biomarkers for prediction and prognosis purposes and how candidate therapeutic targets can be identified.

In CHAPTER 2 a clinical Phase II study is presented, in which the effect of Erlotinib given in front line to patients with advanced NSCLC is studied [76]. Moreover, this chapter describes the correlation of EGFR and KRAS mutations in these patients with response to Erlotinib.

CHAPTER 3 highlights that the correlation of EGFR mutations with response to Erlotinib is not perfect. In this chapter, the case of a Caucasian male, former smoker with squamous cell carcinoma histology and no EGFR mutations, responding to Gefitinib treatment is described [106].
Another case report, described in CHAPTER 4, focuses on three patients presenting with various independent lung lesions. In this chapter it is shown how the use of EGFR and KRAS mutation analysis together with array CGH can help in distinguishing these lesions [107].

Various methods for mutation analysis on both paraffin and frozen samples and using DNA or RNA as template have been described. CHAPTER 5 describes the investigation of possible effects of these different sources of samples on outcome of mutation analysis [108].

In addition to the EGFR mutations, EGFR amplification, detected using Fluorescent In Situ Hybridization (FISH) has been introduced as a biomarker to predict response to EGFR TKIs. In CHAPTER 6 [109], a comparison study of two different techniques to identify EGFR copy number in NSCLC samples Fluorescent- and Chromogenic In Situ Hybridization is described.

Various EGFR mutations have been identified, and it has been shown that sensitivity to EGFR TKIs differs between the different mutations observed. An assay was created to be able to quickly obtain functional information of a specific mutation to the effect on a specific drug. Specific details of this study are described in CHAPTER 7 [110].

Many efforts are ongoing trying to identify new biomarkers for response to EGFR TKIs and much information has been gathered on the EGFR pathway. In order to evaluate the prognostic significance of various EGFR pathway related markers, a combined analysis of mutational and functional activity of various EGFR pathway related molecules in NSCLC patients was carried out, and this study is presented in CHAPTER 8.

In CHAPTER 9, a microarray analysis is described, in which gene expression data is integrated with array CGH information to identify important genes in NSCLC. This chapter focuses on HSP90 whose expression was affected by a deletion on chromosome 14 and was shown to be correlated with survival of early stage NSCLC patients.

Finally, in CHAPTER 10 the findings of all these investigations, as well as future perspectives in this area, are summarized and discussed.
REFERENCES


CHAPTER 2

ERLOTINIB FOR FRONT-LINE TREATMENT OF ADVANCED NON-SMALL-CELL LUNG CANCER. A PHASE II STUDY.

Giuseppe Giaccone, Mariëlle I Gallegos Ruiz, Thierry Le Chevalier, Nick Thatcher, Egbert F Smit, Jose A Rodriguez, Pasi A Janne, Dalila Ould-Aissa and Jean-Charles Soria

CLINICAL CANCER RESEARCH 2006
12: 6049-6055
ABSTRACT

PURPOSE: Erlotinib has proven activity in pretreated patients with advanced non small-cell lung cancer (NSCLC). We evaluated erlotinib in the front-line treatment of advanced NSCLC, and assessed biological predictors of outcome.

EXPERIMENTAL DESIGN: In this phase II study, chemotherapy-naive patients with stage IIIB/IV NSCLC received oral erlotinib, 150 mg/day, until disease progression or unacceptable toxicity occurred. Tumor response was assessed every 6 weeks and samples were analyzed for potential molecular markers of treatment response and survival. The primary endpoint was the proportion of patients without disease progression after 6 weeks of treatment.

RESULTS: 53 patients were eligible. The overall rate of non-progression at 6 weeks was 52.8% (28/53 patients). Tumor response rate was 22.7%, with 1 complete response, 11 partial responses and 16 cases of stable disease. Responses were seen across most patient clinical characteristics. The median duration of tumor response was 333 days, median overall survival was 391 days, and median time to disease progression was 84 days. Erlotinib was well tolerated, the main treatment related adverse events being mild-to-moderate rash and diarrhea. Histological material for biological studies was available in 29 cases. Four out of five responders and one patient with stable disease had a classical EGFR tyrosine kinase mutation. Two progressing patients exhibited EGFR point mutations (one with T790M mutation), and K-ras mutations were detected in 10 non-responders.

CONCLUSIONS: Erlotinib shows significant antitumor activity in the first-line treatment of advanced NSCLC and may be a viable alternative to chemotherapy. Patient selection cannot easily be based on clinical or biological variables.
INTRODUCTION

There are approximately 1.2 million new cases of lung cancer diagnosed every year worldwide, and 1.1 million patients die of the disease [1]. Non-small-cell lung cancer (NSCLC) represents 80% of lung cancers. Surgery is the most important curative modality in the treatment of early-stage NSCLC but, only approximately 20–30% of patients are diagnosed at an operable stage. Platinum-based doublet chemotherapy, the standard treatment for advanced NSCLC, achieves mostly partial responses in only 30–40% of patients and a modest survival increase [2,3] in general single agents are inferior to combinations. Median survival and 1-year survival after first-line chemotherapy are 8–10 months and 30–40%, respectively [4]. These regimens are associated with considerable toxicity, particularly myelosuppression. More effective and less toxic treatments are urgently needed for advanced NSCLC.

Among the new drugs that have been introduced in the treatment of advanced NSCLC after failure of first-line chemotherapy are docetaxel, pemetrexed and, recently, epidermal growth factor receptor (EGFR) tyrosine kinase (TK) inhibitors. All of these agents produce a low response rate (below 10%) and have different toxicity profiles. In a large randomized study, erlotinib improved survival versus best supportive care in second- and third-line treatment [5]. Erlotinib is a targeted agent that inhibits EGFR by competing with ATP at the intracellular TK domain of the receptor. Erlotinib is given orally continuously, and is devoid of bone marrow and major organ toxicities [6]. Common side effects are diarrhea and skin toxicity.

Despite promising preclinical data, the combination of small molecule inhibitors of EGFR with first-line chemotherapy failed to improve survival, compared with 5 chemotherapy alone in four large randomized studies that investigated erlotinib or gefitinib in combination with either carboplatin-paclitaxel or cisplatin-gemcitabine [6].

We performed a phase II study of single-agent erlotinib as first-line therapy in patients with advanced NSCLC. We also performed biological studies on tumor samples of patients entered in the study, to identify patients who benefited most from treatment.
PATIENTS AND METHODS

Patients

Patients with a histological or cytological diagnosis of advanced NSCLC, no amenable to radical surgery or radiotherapy were eligible for the study. No prior chemotherapy or other systemic treatment was allowed. Other eligibility criteria were: measurable disease according to the RECIST criteria, aged ≥18 years, performance status of 0–2, life expectancy > twelve weeks, at least four weeks since any prior surgery or radiotherapy, granulocyte count > 1.5 x 109/L, platelet count > 100 x 109/L, bilirubin and transaminases < 1.5 x upper limit of normal, and serum creatinine < 1.5 x upper limit of normal, or creatinine clearance > 60 mL/min. Females of childbearing potential had to have a negative pregnancy test. Patients were excluded if they had unstable systemic disease (active infection, uncontrolled hypertension, unstable angina, congestive heart failure, myocardial infarction within the previous year, serious cardiac arrhythmia requiring medication), any other malignancies within five years except for adequately treated carcinoma in situ of the cervix, or basal or squamous cell skin cancer. Patients with brain metastases were only allowed in the study if there was no evidence of progression in the brain, and in absence of 6 corticosteroid treatment. Patients with significant eye disorders were excluded (severe dry eye syndrome, Sjogren syndrome, severe exposure keratitis and any other eye disorder likely to increase the risk of corneal epithelial lesions). Signed informed consent was required for all patients. The study was conducted according to the latest version of the Declaration of Helsinki, and the protocol was approved by the independent Ethics Committee and the Review Board from each participating institution.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sense (5’-3’)</th>
<th>Antisense (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR exon 18</td>
<td>CTGACGTTGGGTCCTCTG</td>
<td>TCCGCTACACAGCTGAG</td>
</tr>
<tr>
<td>EGFR exon 19</td>
<td>ATGGACAGCTGTTCTG</td>
<td>GCTGAGACAGCTGAG</td>
</tr>
<tr>
<td>EGFR exon 20</td>
<td>CTGACGTTGGGTCCTCTG</td>
<td>TCCGCTACACAGCTGAG</td>
</tr>
<tr>
<td>EGFR exon 21</td>
<td>TGGACGACAGCTGAG</td>
<td></td>
</tr>
<tr>
<td>PIK3CA exon 9</td>
<td>ATGGACAGCTGTTCTG</td>
<td>GCTGAGACAGCTGAG</td>
</tr>
<tr>
<td>PIK3CA exon 10</td>
<td>TGGACGACAGCTGAG</td>
<td>GCTGAGACAGCTGAG</td>
</tr>
<tr>
<td>KRAS exon 1</td>
<td>AGTCAATTCATGTCATG</td>
<td>TAACTACATTCATGTCATG</td>
</tr>
<tr>
<td>KRAS exon 2</td>
<td>GCCCTGTCATGTCATG</td>
<td></td>
</tr>
<tr>
<td>NR_1</td>
<td>AGTCAATTCATGTCATG</td>
<td>TAACTACATTCATGTCATG</td>
</tr>
</tbody>
</table>

Table 1. Primers used in EGFR, PIK3CA, and KRAS mutation analysis.
CHAPTER 2

Treatment

Erlotinib 150 mg tablets were given orally daily, in the morning with 200 mL of water, and at least one hour before or two hours after ingestion of food or medication. Caution was taken to avoid medications that interact with CYP 450-3A4 enzyme, and concomitant warfarin was not permitted. Assessment of toxicity was according to NCI CTC version 2.0. Dose adjustments, in 50 mg decrements, were performed if grade 3 or 4 toxicity was encountered. Before starting treatment, patients were assessed by physical examination, ECG, disease measurement by appropriate radiological techniques, complete blood cell counts, serum chemistries, and urine pregnancy test in females of childbearing potential. Patients were evaluated every three weeks and hematology and blood chemistry were performed; tumor measurements were assessed every six weeks. An ECG was repeated after eighteen weeks of treatment and as clinically indicated afterwards. Response assessment was according to the RECIST criteria [7]. All patients received the LCSS (Lung Cancer Symptoms Score) patients scale questionnaire at baseline and again at three and six weeks and every six weeks thereafter.

Mutation analysis of EGFR, PIK3CA and K-ras

Paraffin-embedded tumor material was cut into 4 μm-thick sections and placed onto glass slides, stained with hematoxylin and eosin, and the presence of tumor cells was verified by a pathologist. Tumor cells were microdissected and genomic DNA was isolated using the QIAamp DNA MicroKit (Qiagen, Venlo, The Netherlands). Nested PCRs were carried out using primers (Table 1) to amplify exons 18–21 of EGFR, exons 9 and 20 of PIK3CA and exons 1–2 of K-ras. To facilitate sequencing, internal primers incorporated an M13Tag. Sequencing of PCR products was performed with the ABI PRISM™ 310 Genetic analyzer (Applied Biosystems). Mutations were confirmed by sequencing independent PCR products.

Because of concerns about the sensitivity of direct sequencing, DNA from 9 independent samples from Dana Farber Cancer Institute (DFCI) and 13 from the Vrije Universiteit Medical Center were analyzed by the other institution, in a

<table>
<thead>
<tr>
<th>Table 2. Patient characteristics (n=53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Median age (range), y</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Performance status</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>Histology</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
</tr>
<tr>
<td>Bronchioloalveolar carcinoma</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>Smoking History</td>
</tr>
<tr>
<td>Never smoker</td>
</tr>
<tr>
<td>Former or current smoker</td>
</tr>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>IIIB</td>
</tr>
<tr>
<td>IV</td>
</tr>
<tr>
<td>Prior therapy</td>
</tr>
<tr>
<td>Surgery</td>
</tr>
<tr>
<td>Radiotherapy</td>
</tr>
<tr>
<td>Surgery + radiotherapy</td>
</tr>
<tr>
<td>None</td>
</tr>
</tbody>
</table>
blinded fashion, for EGFR mutations in exons 18–21. At DFCI, specimens were analyzed using Surveyor™ DNA endonuclease, combined with the WAVE HS DHPLC system (Transgenomic, Inc) as described [8]. DNA variants detected using Surveyor™, were subjected to fractionation and sequencing.

**EGFR gene copy number**

Chromogenic in situ hybridisation (CISH) was used to study EGFR gene copy number. Tissue sections were incubated at 64°C for 2–4 hrs, deparaffinized and rehydrated. Samples were boiled in 1 mM EDTA/Tris pH 9 and digested with 0.01% 8 pepsin/0.2N HCl at 37°C. After dehydration in alcohol, the EGFR probe was added (Zymed Spotlight EGFR amplification probe 84-1300, Zymed, CA, USA). Samples were denatured at 80°C for 10 minutes and incubated overnight in a humidifying chamber at 37°C. The next day, samples were washed for 5 min in 0.5 x SSC at 75–80°C and incubated in 3% H2O2/PBS for 10 minutes to block endogenous peroxidase activity. After blocking the slides in pre-immune normal goat serum for 10 min, slides were incubated with mouse-α-Digoxigenin antibody (1:100 dilution) for 60 min. CISH signals were visualized using the DAKO Envision™ system (Dakocytomation, Heverlee, Belgium). CISH signals were counted in 200 cells by two observers who were blinded to the patient’s clinical characteristics. The scoring system as described by Cappuzzo et al for FISH [9].

**Statistical considerations**

The primary endpoint of the first part of the study was the rate of nonprogression at 6 weeks, defined as the number of patients whose tumor did not progress (less than 19% tumor increase to total disappearance of tumor, according to RECIST criteria). Deterioration of performance status was considered progression. This end-point was considered an early conservative estimate of biological activity of the compound, and was based on results obtained with single agent first-line chemotherapy [10,11].

![Table 3. Response to erlotinib](image)

<table>
<thead>
<tr>
<th>Response</th>
<th>Patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Partial</td>
<td>11 (21)</td>
</tr>
<tr>
<td>Stable disease</td>
<td>16 (30)</td>
</tr>
<tr>
<td>Progression</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Not evaluable*</td>
<td>13 (25)</td>
</tr>
<tr>
<td>Refusal</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Total</td>
<td>53 (100)</td>
</tr>
</tbody>
</table>

*Less than 6 weeks of treatment.

![Table 4. Treatment-related side effects (n=53)](image)

<table>
<thead>
<tr>
<th>Toxicity</th>
<th>Grade</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td></td>
<td>17 (32%)</td>
<td>28 (53%)</td>
<td>2 (4%)</td>
<td>-</td>
</tr>
<tr>
<td>Diarrhea</td>
<td></td>
<td>30 (57%)</td>
<td>5 (9%)</td>
<td>2 (4%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Hair loss</td>
<td></td>
<td>11 (21%)</td>
<td>5 (9%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nails</td>
<td></td>
<td>3 (6%)</td>
<td>7 (13%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mucositis</td>
<td></td>
<td>7 (13%)</td>
<td>5 (9%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ocular</td>
<td></td>
<td>10 (19%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
<td>-</td>
</tr>
<tr>
<td>Nausea</td>
<td></td>
<td>4 (8%)</td>
<td>8 (15%)</td>
<td>2 (4%)</td>
<td>-</td>
</tr>
<tr>
<td>Vomiting</td>
<td></td>
<td>4 (8%)</td>
<td>2 (4%)</td>
<td>1 (2%)</td>
<td>-</td>
</tr>
<tr>
<td>Transaminases</td>
<td></td>
<td>6 (11%)</td>
<td>3 (5%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bilirubin</td>
<td></td>
<td>4 (8%)</td>
<td>2 (4%)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Sixteen patients were to be entered, and of these at least 8 were to be nonprogressive in order for the study to continue accrual to the second part; if 9 or more patients progressed during the first 6 weeks, the study was to be stopped (Simon’s optimal design; p0=40%, p1=60%, alpha=0.05, beta=0.20). The second part of the study had a targeted total accrual of an additional 30 patients, with a total accrual to 9 the study of 46 patients. If at least 24 patients did not progress, the study would be considered of interest and promising.

Secondary objectives of the study were objective response rate, disease control rate, duration of response, time to disease progression or death, survival, and safety. Patients who did not receive at least 6 weeks of treatment were considered as progressors. Exploratory studies included quality-of-life assessment, and the correlation of biological markers with treatment outcome. Survival, time to progression and duration of response were calculated from day 1 of treatment. Survival curves were constructed using the Kaplan-Meier method, and survival was compared using the Log-rank test. The SPSS statistical package for Windows version 9.01 was used.

### Table 5. Univariate analysis of survival

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median survival (d)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: Male</td>
<td>592</td>
<td>NS</td>
</tr>
<tr>
<td>Female</td>
<td>391</td>
<td></td>
</tr>
<tr>
<td>Stage: IIIB</td>
<td>592</td>
<td>NS</td>
</tr>
<tr>
<td>IV</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td>Performance status: 0</td>
<td>&gt;585</td>
<td>0.077</td>
</tr>
<tr>
<td>1</td>
<td>389</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Smoking status: Never</td>
<td>&gt;644</td>
<td>0.014</td>
</tr>
<tr>
<td>Former/current</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td>Age: &lt;70</td>
<td>592</td>
<td>0.027</td>
</tr>
<tr>
<td>&gt;70</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>Histology: Bronchioloalveolar carcinoma</td>
<td>&gt;607</td>
<td>0.006</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>478</td>
<td></td>
</tr>
<tr>
<td>Squamous</td>
<td>391</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td>Response: CR + PR</td>
<td>627</td>
<td>0.0035</td>
</tr>
<tr>
<td>SD</td>
<td>389</td>
<td></td>
</tr>
<tr>
<td>PD</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>Rash: 0</td>
<td>82</td>
<td>0.0018</td>
</tr>
<tr>
<td>1</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>592</td>
<td></td>
</tr>
<tr>
<td>EGFR: Wild type</td>
<td>377</td>
<td>0.15</td>
</tr>
<tr>
<td>Mutant classic</td>
<td>&gt;627</td>
<td></td>
</tr>
<tr>
<td>KRAS: Wild type</td>
<td>&gt;627</td>
<td>0.0059</td>
</tr>
<tr>
<td>Mutant</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>EGFR-CISH: Amplification + high polysomy</td>
<td>188</td>
<td>0.5</td>
</tr>
<tr>
<td>Other</td>
<td>180</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** CR, complete response; SD, stable disease; NS, not significant; PD, progressive disease; PR, partial response.
RESULTS

From January to July 2004, 54 patients were enrolled in the study from 3 institutions in Europe. One patient was not eligible because of unstable brain metastases, and did not start treatment. Table 2 summarizes the main characteristics of the eligible patients.

Response

Of 16 patients entered into the first part of the study, 10 had not progressed after 6 week's treatment; therefore, the study was opened to further accrual. Thirteen patients had less than 6 weeks' treatment; 12 were classified as early progressions and 1 refused further treatment. The response rate based on intent to treat was 22.7% (95% confidence intervals (CI) 12.3–36.2%) (Table 3). The percentage of nonprogressors at 6 weeks was 52.8% (28/53; 95% CI 38.6–66.7%). The median duration of response was 333 days.

Responses were significantly correlated with histology (6 responses in adenocarcinomas, 4 in bronchioloalveolar carcinomas (BAC), 2 in squamous carcinomas, and none in large cell and other histologies; Pearson Chi square \( P = .025 \)), smoking history (7 responses in never-smokers and 5 in former/current smokers; Pearson Chi square \( P = .016 \)), but not gender, performance status, stage of disease, or age.

When patients progressed, they were offered chemotherapy. Chemotherapy was administered to 31 patients, 28 of whom received carboplatin or cisplatin in combination with gemcitabine; the response rate was 28.6% in these patients. Three patients received single-agent chemotherapy, and 1 patient was rechallenged with 11 erlotinib after a partial response and interruption of treatment for 8 months at the patient's request; this patient again experienced a partial response. No further therapy was given to 13 patients who progressed on first-line erlotinib (24.5%), and 4 had only palliative radiotherapy. The median survival of progressive disease patients was 234 days, versus 592 days for those with disease control (\( P = .004 \)). The 13 patients who did not receive any form of treatment at progression had a significantly worse performance status at start (Likelihood ratio Chi square \( P = .020 \)). The median survival of the 28 patients who received platinum-based chemotherapy was 350 days, calculated from the date of progression on erlotinib.

Toxicity

The most frequent drug-related side effects were mild to moderate skin toxicity and diarrhea (Table 4). All responders had some degree of rash, whereas only 7 nonresponders developed rash (Likelihood ratio Chi square \( P = .048 \)). Emesis was usually mild. Liver toxicity (transaminase and bilirubin increase) was observed in just over 10% of patients and was of mild or moderate
severity and reversible. Dose reduction to 100 mg/day was necessary in 5 cases due to increased bilirubin, and in 1 due to skin toxicity. Dose interruptions for more than one day were necessary in 13 patients; in 5 this was due to hyperbilirubinemia.

<table>
<thead>
<tr>
<th>Table 6. Mutation analysis and patient characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Gender Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Histology Adenocarcinoma</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>Smoking status Never-smoker</td>
</tr>
<tr>
<td>Current/former smoker</td>
</tr>
<tr>
<td>Response Partial response</td>
</tr>
<tr>
<td>Stable disease</td>
</tr>
<tr>
<td>Progressive disease</td>
</tr>
</tbody>
</table>

*Only detectable by WAVE.

Quality of life

The LCSS was analyzed by assessing the visual analogue averages to the first questions, concerning disease related symptoms (appetite, fatigue, coughing, shortness of breath, hemoptysis, pain). A decrease of 25% of the score compared to basal level was considered symptomatic progression. Although there was a difference in time to symptomatic progression between responding patients, stable disease and progressing patients, this was not significant (not shown).

**EGFR, K-ras and PIK3CA mutations and EGFR copy number**

Histological material from 29 patients was available to perform mutation analyses. EGFR mutations were detected in 7 cases, 5 of which were deletions in exon 19; 4 of these cases responded to therapy and 1 had long-lasting stable disease. One major response was detected in a patient with wild-type EGFR. Furthermore, 2 point mutations were detected in exon 20, one not previously described (V802I) and one that has been described in resistant disease (T790M) [12,13]. In the latter sample, another point mutation in exon 21 was also recognized (R832H). These point mutations, which were only detectable using the WAVE system, but not direct sequencing, were observed in progressing patients and in tumors also bearing a K-ras mutation.
The exon 19 deletions were only observed in females (Fisher’s Exact test $P = .047$), mainly in never-smokers (Fisher’s Exact test $P = .030$). No patients with K-ras mutations had an exon 19 deletion (Likelihood ratio $P = .026$). There was no association between classical EGFR mutations and histology, performance status, stage, age, or appearance of rash upon treatment.

K-ras mutations were detected in 10 patients, all smokers (Fisher’s Exact test $P = .061$), and non-responders. There was a trend for a correlation between K-ras mutations and lack of development of rash (Fisher’s Exact test $P = .052$), and no association with gender, histology (though more frequent in adenocarcinomas), performance status, stage, or age.

The presence of classical EGFR mutations was significantly correlated with response to treatment (Fisher’s Exact test $P = .001$), and K-ras mutations were more commonly associated with lack of response (Fisher’s Exact test $P = .125$). One PIK3CA mutation was detected in a patient who had stable disease on treatment. A summary of the mutation analyses is reported in Table 6. Survival was longer in patients with classical EGFR mutations than in the other patients, although the difference did not reach statistical significance (Table 5 and Fig. 1A). Conversely, survival of patients with K-ras mutations was significantly shorter than that of patients with wild-type K-ras (Table 5 and Fig. 1B). Of the two patients still on treatment, one did not have tumor available for mutation analysis; the other patient’s tumor contained an exon 19 deletion of EGFR.

EGFR CISH analysis could only be performed in 13 tumors, due to insufficient tumor tissue remaining after mutation analyses. Amplification was detected in 2 cases, and high polysomy in 4. One of the 2 samples with amplified EGFR also contained an exon 19 deletion, whereas the other sample contained a K-ras mutation. No clear correlation was observed between EGFR copy number and response to treatment (Fisher’s Exact test $P = .192$) or survival.
DISCUSSION

EGFR TK inhibitors are active drugs in advanced NSCLC, but in unselected patients the response rate in second- and third-line for both erlotinib and gefitinib is in the range of 10% or lower [6]. However, response to EGFR TK inhibitors is higher in women, never-smokers, East Asians, adenocarcinoma and bronchioloalveolar carcinomas, and those who develop rash [14]. Furthermore, activating mutations in the EGFR TK domain, and EGFR amplification have been shown to be predictors of response and survival in several reports [9,15-17].

In the present study the primary endpoint was met, with more than 50% of patients showing no progression after 6 weeks’ treatment. The response rate of 22.7% is similar to that obtained with most single chemotherapy agents. Although the protocol did not prescribe any type of patient selection, a higher percentage of women, never-smokers, and patients with BAC histology were included in the study than one would expect in a typical sample of patients with advanced NSCLC. This was a result of the growing evidence, accumulating during the conduct of the study, that these characteristics are more often associated with benefit from EGFR TK inhibitors. Although response to erlotinib was more common in patients with adenocarcinoma (including BAC histology) and in never-smokers, like shown in other studies [5,18], responses were also observed in patients with tumors of other histology, and in former/current smokers. For example, a complete response was reported in a male patient who was a smoker. These results are consistent with those observed in the BR.21 study, in which erlotinib produced a significant survival benefit compared with best supportive care in most of the patient subgroups analyzed [5].

In the present study, the median overall survival (approximately 13 months) and the 1-year survival rate of 54% were better than those reported for several platinum-based doublet chemotherapy regimens. The use of erlotinib in first-line did not appear to have a detrimental effect on the response to subsequent chemotherapy. Our results are similar to those obtained with first-line gefitinib in Japanese patients [19]. Promising results with erlotinib 150 mg/day monotherapy have also been achieved in an ongoing phase II study of patients aged over 70 years with previously untreated advanced NSCLC [20].

Erlotinib was well tolerated in this study with the main toxicities being rash and diarrhea. Both response and survival were correlated with the grade of rash, in agreement with several other trials [18].

Several studies have assessed integration of EGFR inhibitors into first-line treatment of advanced NSCLC. Although the concomitant administration of erlotinib [21,22] and gefitinib [23,24] with chemotherapy failed to improve survival in the overall population, a survival benefit was observed in never-smokers who received erlotinib and chemotherapy [25]. This suggests that patient
selection may identify groups of patients who may derive a large benefit from erlotinib treatment in combination with chemotherapy.

Certain mutations in the TK domain of EGFR are associated with increased sensitivity to EGFR TK inhibitors [16,17,26]. In our study, four out of five responding patients had EGFR mutations represented by well-documented deletions in exon 19. However, the tumor of one responding patient did not possess an EGFR mutation, and a mutation was found in a patient with long-lasting stable disease. In addition to exon 19 deletions, 2 point mutations were identified in patients who progressed, including the known resistant mutant T790M and a novel point mutation. Interestingly, in these samples K-ras mutations were also found, suggesting that some EGFR point mutations can be associated with K-ras mutations and are related to resistance. The T790M EGFR mutation has been described in patients who were initially sensitive to EGFR TK inhibitors as a result of an exon 19 deletion, and who later became resistant [12,13]. However the T790M mutation has also been described in a patient with primary resistance [27], and germline inheritance of this variant has been proposed in one family [28]. The use of irreversible EGFR TK inhibitors or monoclonal antibodies directed to the extracellular domain of EGFR might circumvent resistance in these cases [12,29-32]. In our study, the presence of classical EGFR mutations was correlated with longer survival, although, possibly due to the small number of samples analyzed, this did not reach statistical significance. We also analyzed the EGFR copy number using CISH, but were unable to confirm the predictive value of this parameter in our series, although the number of samples analyzed was too small to draw any firm conclusion.

In the BR.21 study, several techniques were utilized to assess EGFR status: mutation analysis, FISH and immunohistochemistry [33], and FISH appeared to be the strongest predictor of response and survival, although by multivariate analysis none of the techniques were able to significantly predict outcome. In a retrospective analysis of patients treated with gefitinib, alone or in combination with chemotherapy, EGFR amplification (detected by a PCR method) and EGFR mutations identified two distinct groups of patients, and mutations were better predictors of outcome than amplification [34].

Technological issues are still very important for both the detection of mutations and the assessment of EGFR copy number [14]. In a cross-validation analysis (data not shown) it was apparent that simple direct sequencing does not provide sufficient sensitivity for detection of all mutations; however, the use of more sensitive methods may identify mutations of uncertain significance [8,14]. Tumor heterogeneity and possibly the presence of mutations in a proportion of tumor cells may in fact potentially play a role in the level of detectability and biological relevance.

K-ras mutations appear to be associated with primary resistance to EGFR TK inhibitors [35]. In our study, K-ras mutations were noted only in patients who did not respond to erlotinib, and
survival was also significantly shorter. Poorer clinical outcomes were shown also for K-ras mutant patients treated with erlotinib plus chemotherapy [36], suggesting that K-ras mutations not only confer a poor prognosis, but may also constitute a negative factor in patients treated with EGFR TK inhibitors. Interestingly, the addition of erlotinib to gemcitabine chemotherapy has been shown to significantly increase survival in patients with advanced pancreatic cancer [37], a tumor type in which K-ras mutations are very common, raising the question of whether the negative findings in NSCLC are limited to this tumor type. Potentially, the simultaneous inhibition of EGFR and ras signaling may be of interest in NSCLC [38].

The results of several studies show that EGFR mutations and K-ras mutations are mutually exclusive [39,40]. Interestingly, in our study classical EGFR mutations and K-ras mutations were mutually exclusive; however, tumors from 2 refractory patients possessed both EGFR point mutations and K-ras mutations, indicating that the mutual exclusivity may only apply to activating mutations of EGFR. In the attempt to select patients for EGFR TKI, the exclusion of patients with K-ras mutations should be considered.

Given the difficulty of combining erlotinib and gefitinib with concomitant chemotherapy in first-line, other strategies need to be developed. The promising efficacy of erlotinib monotherapy in this study suggests that such therapy is a potential option for the first-line treatment of advanced NSCLC. Studies investigating the optimal sequence of erlotinib and chemotherapy are warranted. Selection of patients based on clinical and biological characteristics will allow a significant improvement of overall results. Combinations of targeted therapies also warrant investigation in this setting: erlotinib plus bevacizumab produced promising antitumor activity in a phase I/II trial involving patients with previously-treated advanced NSCLC [41], and this combination is now also being investigated in first-line.

In summary, the results of this study show that erlotinib has significant antitumor activity in the first-line treatment of advanced NSCLC, which may be comparable with single agent chemotherapy, and is devoid of significant side effects. However, these results have to be seen in the context of the patient population entered, which is weighted toward individuals with higher chances of having EGFR mutations. Although responses to erlotinib were more common in some subgroups of patients, clinical benefit was not restricted to these subgroups. Further investigation of erlotinib monotherapy in this setting is warranted and the identification of patients who most benefit from the treatment is of paramount importance.
REFERENCES


ACKNOWLEDGMENTS

We would like to thank Dr Gerrit Meijer and Karijn Floor for the CISH analysis and Alison J. Holmes for the cross-validation of EGFR mutations. The authors would also like to thank Dr M Smith for his assistance in drafting the manuscript. Dr Smith is a medical writer with Thomson Gardiner-Caldwell Communications.
CHAPTER 3

RESPONSE TO ERLOTINIB IN THE FIRST-LINE TREATMENT OF NON-SMALL CELL LUNG CANCER IN A CAUCASIAN MALE SMOKER WITH SQUAMOUS-CELL HISTOLOGY

Mihaela Achille, Mariëlle I Gallegos Ruiz, Giuseppe Giaccone and Jean-Charles Soria

CLINICAL LUNG CANCER 2006
8: 214-216
ABSTRACT

Erlotinib, a potent inhibitor of the tyrosine-kinase (TK) activity of human epidermal growth factor receptor (HER1/EGFR), produces significant survival, symptom and quality-of-life benefits in previously treated advanced non-small-cell lung cancer (NSCLC). The survival benefit from erlotinib is observed in a broad spectrum of patients, although radiographic responses are observed more commonly in certain patient subgroups (females, never smokers, adenocarcinoma histology, Asian ethnic origin, presence of HER1/EGFR TK domain mutations). Here we describe a good response to first-line erlotinib in a Caucasian male smoker with advanced squamous-cell NSCLC. A molecular analysis showed that the tumor in this patient did not possess HER1/EGFR TK mutations. This case study, along with subgroup analyses of the BR.21 phase III study, suggests that patients should not be selected for erlotinib treatment based only on characteristics such as smoking status, tumor histology, HER1/EGFR TK mutation status, gender or ethnic origin.
CASE REPORT

A 55-year old Caucasian male patient was diagnosed with non-small-cell lung cancer (NSCLC) after presenting with supraclavicular lymph node enlargement in March 2004. The patient had a history of chronic smoking (10 to 20 cigarettes per day for 30 years; stopped smoking at diagnosis) and drinking (stopped at diagnosis). Chest radiography and computer tomography (CT) scans revealed a 55 mm mass in the right hilum (Figure 1A), with mediastinal and right supraclavicular nodal enlargement. The size of the left adrenal gland was increased on CT scan, suggesting the possibility of distant metastasis (ct2N3Mx, stage IIIB or IV). The patient’s Eastern Cooperative Oncology Group (ECOG) performance status was 1 and laboratory findings were normal. Fibrobronchoscopy revealed a tumor occluding the right upper lobar bronchus and tumor tissue was obtained via biopsy. Histopathology revealed squamous-cell carcinoma, and immunohistochemistry (IHC; using a 0 to 3+ scale based on staining intensity) showed strong positive (3+) membrane staining for human epidermal growth factor receptor (HER1/EGFR). IHC analyses were also performed for downstream signaling molecules. Positive staining was observed for phospho-Akt in both the nucleus (3+) and the cytoplasm (2+). Weak cytoplasmic staining (1+) and a lack of nuclear staining was observed for both phospho-Erk and phospho-STAT5. No mutations were observed in exons 19, 20 or 21 of the HER1/EGFR gene, which encode part of the tyrosine kinase domain; other exons were not sequenced because of insufficient material. Fluorescence-in situ hybridization analysis for HER1/EGFR gene amplification was not performed because tissue sample from the biopsy was exhausted.

![Figure 1](image-url) (A) CT scans before treatment and (B) after 39 weeks of erlotinib therapy (21 weeks at 150 mg/day, and 18 weeks at 100 mg/day)
ERLOTINIB IN NSCLC OF SQUAMOUS-CELL HISTOLOGY

TREATMENT

Treatment with erlotinib 150 mg/day was initiated on 28 April 2004 as part of a phase II clinical trial evaluating the first-line treatment of advanced NSCLC with erlotinib monotherapy. After 6 weeks of treatment, a CT scan revealed a very good partial response: the right hilar mass had decreased to 30 mm. The supraclavicular lymph nodes disappeared and the patient was asymptomatic. After an additional 6 weeks of erlotinib treatment, the right hilar mass had further reduced to 20 mm. Bilirubin levels assessed on 11 August 2004 were elevated (23 μmol/L), resulting in interruption of erlotinib treatment for 1 week until bilirubin levels were within the normal range (0-17 μmol/L). Bilirubin levels were elevated again on 8 September 2004 (28 μmol/L). After an additional interruption of erlotinib treatment for a period of 1 week, the patient resumed treatment with erlotinib at a lower dose of 100 mg/day.

After 24 weeks of erlotinib treatment, the patient’s right upper lobe mass and left adrenal gland were stable, and the mediastinal lymph nodes were less than 1 cm. At this time bilirubin levels were elevated and an additional 1 week suspension of erlotinib treatment was necessary. Erlotinib was well tolerated throughout the treatment course, with only minimal skin toxicities (limited to grade 1 with few pustular elements and 2 episodes of grade 2 eruption). Skin toxicities were readily managed using topical clindamycin (twice daily) and moisturising creams.

After 30 weeks of treatment with erlotinib, a positron emission tomography (PET) scan revealed high uptake of 18F-fluorodeoxyglucose (FDG) by the right hilar mass. No uptake of FDG by the mediastinal lymph node or left adrenal gland was observed. A CT scan at 36 weeks revealed an increase in the right upper lobe mass from 20 to 30 mm. At this time, the patient agreed to surgical resection. A total of 39 weeks of erlotinib therapy was administered prior to surgery (Figure 1B). On 28 February 2005 the patient underwent a lobectomy and mediastinal lymph node dissection. Histopathology showed a 35 mm moderately differentiated epidermoid carcinoma with 10% necrosis, a clear surgical margin, involvement of 2 peribronchial lymph nodes and no mediastinal lymph node invasion (pT2N1).

Four cycles of adjuvant chemotherapy with cisplatin plus gemcitabine (1 cycle) and carboplatin plus gemcitabine (3 cycles) were administered from 12 April to 10 June 2005. During the first course of chemotherapy, the patient experienced dizziness, paraesthesia and muscular weakness of the lower limbs. A brain CT scan and magnetic resonance imaging (MRI) showed an ischemic stroke caused by diffuse atherosclerosis. After this event, the patient presented with occasional muscular weakness of the left lower limb. As assessed by PET and CT scans, the patient has now remained free from NSCLC for the past 10 months.

Erlotinib, a HER1/EGFR tyrosine-kinase inhibitor (TKI), is indicated in NSCLC for the treatment of patients with advanced disease who have failed at least one prior chemotherapy regimen. US
Food and Drug Administration approval for this indication was based on the results of a well-designed, placebo-controlled phase III study (BR.21), conducted by the National Cancer Institute of Canada-Clinical Trials Group, in which erlotinib 150 mg/day demonstrated a clear survival benefit as second- or third-line treatment in patients with stage IIIIB or IV NSCLC [1]. In this study, erlotinib significantly prolonged both overall survival (6.7 vs. 4.7 months; hazard ratio [HR] 0.70, \( P < 0.001 \)) and progression-free survival (2.2 vs. 1.8 months; HR 0.61, \( P < 0.001 \)) relative to best supportive care. In contrast, the results of a similarly designed study, Iressa Survival Evaluation in Lung Cancer (ISEL), failed to demonstrate a statistically significant survival advantage from the HER1/EGFR TKI gefitinib compared with best supportive care, either in the overall population (overall survival 5.6 vs. 5.1 months; HR 0.89, \( P = 0.087 \)), or in patients with adenocarcinoma (overall survival 6.3 vs. 5.4 months; HR 0.84, \( P = 0.089 \)) [2].

**DISCUSSION**

There is considerable interest in determining predictive markers to identify and select patients most likely to respond to HER1/EGFR-targeted agents, with the aim of improving patient outcomes [3]. Although response rates to erlotinib were higher in certain subgroups of patients in BR.21, exploratory analyses from that trial revealed that erlotinib had a beneficial survival effect in almost all subgroups, regardless of performance status, gender, tumor histology, smoking status, age, HER1/EGFR expression, or race/ethnic group [1]. Reports of increased responsiveness to HER1/EGFR TKIs in patients with NSCLC harboring gain-of-function mutations in the HER1/EGFR gene [4] or in patients with increased HER1/EGFR gene copy number [5] have led to speculation that HER1/EGFR mutation status or copy number could be used as a clinical predictive marker. However, multivariate analyses of tumor biopsy samples collected from patients enrolled in BR.21 determined that whereas expression of HER1/EGFR, an increased
number of HER1/EGFR gene copies, and the presence of HER1/EGFR TK mutations were associated with increased tumor response rates to erlotinib (significant for HER1/EGFR gene copy number and the presence of HER1/EGFR TK mutations), the survival benefit from erlotinib therapy was not influenced by any of these factors [6,7].

The present case study highlights the significant antitumor activity of erlotinib administered in a phase II trial as first-line monotherapy to a Caucasian male patient with HER1/EGFR wild-type squamous-cell carcinoma who smoked up to the time of diagnosis. In this patient, after 39 weeks of erlotinib treatment, the tumor was surgically resectable and the patient remains free from disease 10 months later. When administered as first-line monotherapy, erlotinib was well tolerated with mild-to-moderate skin rash reported in this patient. This is reflective of the overall tolerability profile in the phase II study, where mild-to-moderate easily-managed skin rash and diarrhea were the most common treatment-related adverse events [8]. In this phase II study, patients with adenocarcinoma (including bronchioloalveolar-carcinoma histology) and nonsmokers were more responsive to erlotinib monotherapy; however, treatment responses were also observed in patients with other histology and in current and former smokers. Furthermore, a retrospective analysis of the BR.21 study showed a significant survival benefit from erlotinib in the group of male, ever-smoking patients with squamous-cell carcinoma (median survival 5.5 months vs. 3.4 months with placebo; HR 0.66, \( P = 0.016 \)) [9]. Collectively, the results of this case study along with the subgroup and HER1/EGFR analyses from BR.21 suggest that selection of patients for erlotinib therapy should not be based solely on HER1/EGFR status or patient characteristics such as tumor histology, smoking history, gender, and ethnic origin.
REFERENCES


ACKNOWLEDGMENTS

The authors would like to thank Dr Mark Smith of Gardiner-Caldwell Communications for his assistance in drafting the manuscript. This support was funded by F. Hoffmann-La Roche Ltd.
CHAPTER 4

GENETIC HETEROGENEITY IN PATIENTS WITH MULTIPLE NEOPLASTIC LUNG LESIONS; A REPORT OF THREE CASES

Mariëlle I Gallegos Ruiz, Hester van Cruijsen, Egbert F Smit, Katrien Grünberg,; Gerrit A Meijer, José A Rodriguez, Bauke Ylstra and Giuseppe Giaccone

JOURNAL OF THORACIC ONCOLOGY 2007
2: 12-21
ABSTRACT

Introduction: In patients presenting with multiple malignant lung tumors, determining the relation between the various lesions is important in order to define the best treatment approach. A better understanding of the molecular alterations present in the different lesions may help in defining this relation.

Methods: We carried out a detailed molecular analysis of several tumor specimens obtained from three patients presenting with multiple lung lesions. Tumor specimens were analyzed for EGFR and KRAS mutations by direct DNA sequencing. In addition, a genome-wide chromosomal copy number analysis was performed on DNA extracted from the various lesions using array-based Comparative Genomic Hybridization.

Results: In one case, a deletion of 15 base pairs in exon 19 of EGFR was present in all tumor sites analyzed. Furthermore, a similar pattern of chromosomal aberrations was observed among the various lesions, suggesting that they share the same clonal origin. In the other two cases, in contrast, we identified distinct KRAS genotypes among the various lesions from the same patient. These lesions, moreover, showed different chromosomal aberration patterns, indicating that they may have different underlying pathways of tumorigenesis.

Conclusion: Our results show that EGFR and KRAS mutation analysis, combined with chromosomal copy number profiling can help in defining the relationship between different tumors in one patient.
INTRODUCTION

Non-small cell lung cancer (NSCLC) is the leading cause of death from cancer in the United States and in Western Europe. Most patients with NSCLC are candidates for systemic therapy at some point in the course of the disease.

Particular molecular tumor characteristics can be of help in deciding on systemic treatment. For example, certain mutations in the epidermal growth factor receptor (EGFR) and KRAS genes have been shown to correlate with either favorable or unfavorable patient response, respectively, to treatment with the EGFR tyrosine kinase inhibitors (TKI) gefitinib and erlotinib [1-3].

NSCLC patients are at relatively high risk of developing secondary primary lung tumors, depending on stage at initial diagnosis, with percentages ranging from 3-14% [4-6]. Lungs are moreover a major site of metastases from common solid tumors. In patients with multiple lung tumors, the clonal relation of the various tumors is important in deciding on treatment.

Histopathological examinations of the multiple lesions are often not conclusive as to the differential diagnosis of second primary or metastasis, and a more extensive examination of the genomic changes that occur may help in this diagnosis. In this regard, microarray based Comparative Genomic Hybridization (array-CGH), that measures DNA copy number changes throughout the entire genome can provide important diagnostic information [7-9].

We performed a detailed genetic analysis of tumor samples obtained from three patients with multiple malignant lung lesions using EGFR and KRAS mutation analysis combined with array CGH, which allowed identification of clonal relationships among the various lung lesions.

CASE REPORTS

Patient 1

In April 2003, a 29 year-old male presented to his primary care physician with thoracic pain and hemoptysis. He had a performance status of 0 and showed no weight loss. Physical examination showed no abnormalities. Besides leucocytosis (12·4x10^9/L), laboratory values were normal. The patient had no significant medical history and used no medication. He worked as a professional mason, and was a very light smoker having a smoking history of 1 pack-year.

A CT scan of the thorax showed a right lower lobe mass. Subsequently, a diagnosis of pneumonia was made but despite treatment with antibiotics, no clinical or radiographic improvement was observed. Because of persistent hemoptysis a bronchoscopy was performed that showed no endobronchial lesions. The CT scan of the thorax was repeated and revealed the known lesion in the right lower lobe and two enlarged mediastinal lymph nodes, ie: mediastinal nodal station 4R.
FIGURE 1. Localization and molecular analysis of tumors from patient 1. (A) A schematic illustration of localization and time of diagnosis of multiple malignant lung lesions in patient 1. (B) CT scans of the thorax show lesion 1 in the right lung before and after chemoradiation, and lesion 2 and 3 in the left lung. (C) Genome wide profiles of tumor sites 2 and 3. Tumor DNA of patient 1 was hybridized with normal female reference DNA on a 30K human oligonucleotide array. On the X-axis, array elements are ordered according to their position on the genome. Odd and even chromosomes are indicated in light grey and black, respectively. On the Y-axis ratios are plotted as log2 ratio of the fluorescent intensities of tumor (Cy3) divided by reference (Cy5) signal for each element on the array. Smoothed values are indicated in dark grey. A smoothed log2 ratio close to zero indicates no difference in fluorescence intensity between tumor and reference DNA and hence, no chromosomal copy number aberrations. Log2 ratios higher or lower than zero indicate gains or losses of chromosomal elements, respectively. For patient 1, similar patterns of chromosomal aberrations can be observed for tumor sites 2 and 3, indicating that these lesions share the same clonal origin.
and 7. The lesion in the right lower lobe and mediastinal lymph nodes had uptake on an $^{18}$FDG PET scan. Cytological analysis of the subcarinal lymph node, obtained by a EUS-FNA showed adenocarcinoma. The tumor was staged as IIIaN2 non-small cell lung carcinoma (T2N2M0). The patient was treated on a chemoradiation protocol that employed platinum based chemotherapy and definite thoracic radiotherapy. After the end of treatment, a CT scan and an $^{18}$FDG PET scan were repeated. CT scan of the thorax showed a partial response of the lesion (Figure 1B). $^{18}$FDG PET scan showed no uptake in the primary lesion, and minimal uptake in a subcarinal lymph node. Biopsies of mediastinal lymph nodes obtained by mediastinoscopy showed no malignant cells, and the patient underwent an R0 bilobectomy. Histological analysis revealed residual adenocarcinoma at the primary tumor site and no detectable tumor in hilar and mediastinal lymph nodes. The patient recovered well after surgery. No adjuvant therapy was administered.
In February 2005, the patient reported again slight hemoptysis. A CT scan showed two small lesions in the left lung; one in the left lower lobe and one in the left upper lobe. These two lesions, as well as two left hilar lymph nodes, were positive on an [18FDG] PET scan. The lesion in the left lower lobe appeared radiologically compatible with a second primary tumor, and in retrospect actually was already visible on the CT scans performed before the operation in 2003. At that time this was not considered to be tumor, because of the lack of FDG avidity and the lack of change after induction chemotherapy (Figure 1B). The patient underwent a segmental resection of the left upper lobe, a segmental resection of the left lower lobe and mediastinal lymph node dissection (stations 4, 5, 8 and 10) in May 2005. Histological analysis showed poorly differentiated adenocarcinoma in both lung nodules and two mediastinal lymph nodes (5 and 10) contained metastatic adenocarcinoma. Postoperatively he received three cycles of cisplatin and vinorelbine. A tyrosine kinase inhibitor of EGFR was administered after the end of chemotherapy. The patient is well without evidence of disease after more than one year since the operation.

Patient 2

In May 2005, a 48-year-old woman was referred by her primary care physician because of persisting cough. An X-ray of the chest showed two contralateral pulmonary lesions. She was in a good clinical condition and had no weight loss. Physical examination revealed no abnormalities. Besides a mild elevation of the erythrosedimentation rate (27 mm) and leukocytosis (13.3x10^9/L), laboratory values were normal. She had a smoking history of 35 pack-year, and used fluoxetine for depression. Her medical and family history were negative for malignancy.

A CT-scan of the thorax showed a cavitating lesion in the right upper lobe and a solid lesion in the left lower lobe (Figure 2B). There were no mediastinal lymphadenopathies. Pulmonary function tests and diffusion capacity were normal. Bronchoscopy showed no endobronchial lesions; cytologic evaluation of a broncho-alveolar lavage of the left lower lobe identified adenocarcinoma cells. An [18FDG] PET-scan showed the two pulmonary lesions and a hot spot in the right lobe of the thyroid gland. Histological analysis of the thyroid lesion revealed a papillary thyroid carcinoma.

In June 2005, the patient underwent complete resection of the thyroid gland and resection of the cervical lymph nodes. On histopathological evaluation of the resection specimens a papillary thyroid carcinoma of 8 mm was found while all resected lymph nodes were negative for malignancy.

Since the post-operative 131I-scanning showed no uptake in the pulmonary nodules and the early stage of the papillary thyroid carcinoma, a tentative diagnosis of double tumor was made: adenocarcinoma of the lung and papillary thyroid carcinoma.
In August 2005, she was referred to our hospital and an $^{18}$FDG PET-scan showed only uptake in both pulmonary lesions. In September 2005, a wedge resection of the tumor in the right upper lobe was performed. Pathological analysis showed moderately differentiated adenocarcinoma of the lung. Immunohistochemical staining for thyroglobulin was negative. In October 2005, the lower lobe of the left lung and lymph nodes 7 to 11 were also resected. Pathology showed again a moderately differentiated adenocarcinoma of the lung with no positive staining for thyroglobulin. All resected lymph nodes appeared free of malignancy. Patient recovered well from both resections. Five months after the latter operation the patient presented with neurological symptoms and brain metastases were diagnosed. She received whole brain radiotherapy and eventually died of her disease.

**FIGURE 2.** Localization and molecular analysis of tumors from patient 2. (A) Schematic illustration of localization and time of diagnosis of the various malignant lung lesions of patient 2. (B) CT scans of the thorax. Lesion 2 can be observed in the right lung and lesion 3 in the left lung. (C) Chromosomal aberration plots are depicted of lesion 2 and 3. Tumor DNA of patient 2 was hybridized with normal male DNA on a 6K human BAC array. Log2 ratios are displayed against their position on the genome. Odd and even chromosomes are indicated in light grey and black, respectively. Smoothed values are indicated in dark grey. The chromosomal aberration patterns differ between tumor site 2 and 3. In particular, site 3 shows a log2 ratio of -0.35 at chromosome 3, 6, 9, 10, 14q (arrows), these areas have a lower intensity of fluorescence in tumor DNA versus reference DNA, indicating that these chromosomal regions are deleted. The smoothed log2 ratio of +1.4 on chromosome 14p (arrow) indicates that this chromosomal area has a high copy number in comparison to the reference DNA. These narrow, high level gains are referred to as amplification. Thus in patient 2, less chromosomal aberrations are present in site 2 versus site 3, indicating they arose from different clones.
Patient 3

A 65-year-old woman with a 50 pack-year history of smoking was seen by her family physician in February 1997 because of mild dyspnea on exertion, asthenia and fatigue.

An X-ray of the chest revealed a 3 cm lesion in the right lower lobe, which was confirmed on CT-scan of the thorax. The lesion was not radiologically typical for malignancy and there were no lymphadenopathies. A CT-scan after 5 months showed no increase in the dimension of the lesion. An $^{18}$FDG PET scan showed minimal uptake in the lesion in the right lower lobe, but increased uptake in several mediastinal lymph nodes. No endobronchial lesions were seen on bronchoscopy and cytology of broncho-alveolar lavage revealed no malignant cells. Lymph node biopsies
obtained by mediastinoscopy showed no malignant cells. A thoracotomy was performed in March 1998 and the lesion in the right lower lobe was resected. The intra-operative pathological analysis revealed papillary adenocarcinoma, and the surgical procedure was limited to a lobectomy of the right lower lobe (pT2N0M0).

Post-operatively, the patient experienced continuous thoracic pain, which was resistant to several pain medications. Entrapment of an intercostal nerve due to callus formation after the surgical procedure was hypothesized as a potential cause for her pain, which might benefit of a partial resection of ribs 6 and 7. A new CT-scan of the thorax showed two subpleural lesions in the right lung in November 2002. One of these two lesions was positive on a [18FDG] PET scan. A partial resection of ribs 6 and 7 was performed, and a wedge resection of the two aforementioned lesions in the right lung was carried out in March 2003. Pathological analysis showed papillary adenocarcinoma of the lung, which was radically resected. No adjuvant therapy was given. Unfortunately, the resection of ribs 6 and 7 did not have the desired palliative effect on her thoracic pain. A CT-scan performed in January 2006 showed multiple intrapulmonary lesions in the right and left lung. Pathological examination of a CT-guided biopsy confirmed papillary adenocarcinoma. The patient consented to an investigational treatment with bevacizumab and erlotinib. After 4 courses of therapy she had signs of disease progression and protocol treatment was halted. She is currently alive with disease.

MATERIALS AND METHODS

Samples

Tumor specimens obtained prior to therapy were collected, informed consent was obtained from all patients. Paraffin embedded tissue blocks were cut into 4 μm-thick sections and placed onto glass slides. Slides were stained with hematoxylin and eosin and a pathologist (GAM) verified the presence of tumor cells.

Mutation analysis of EGFR and KRAS genes

Isolation of genomic DNA and nested PCRs were carried out to amplify exon 18-21 of EGFR and exon 1-2 of KRAS, as previously described [10]. Sequencing of PCR products was performed using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA), with the ABI PRISM™ 3100 Genetic analyzer (Applied Biosystems). Mutations were confirmed by sequencing independent PCR products. In doubtful cases the PCR product was cloned into the pGEM®-T easy Vector (Promega Benelux BV, Leiden, The Netherlands) and several independent clones were sequenced.
Array Comparative Genomic Hybridisation

DNA isolation from paraffin embedded material was carried out as previously described [11]. DNA labeling and hybridization on CGH BAC and oligonucleotide microarrays was performed as described by van den Ijssel et al [12]. In short, genomic DNA (300 ng) was labeled with Cy3-dCTP (tumor sample) or Cy5-dCTP (reference sample of the opposite sex) (Perkin Elmer, Boston, MA, USA). Labeled, precipitated DNA was dissolved in hybridization mixture and hybridized to the 30K human oligo (patient 1 and 3) or to a 6K human BAC (patient 2) array using a GeneTAC/HybArray12 hybstation (Genomic Solutions, Perkin Elmer). Slides were scanned using a Microarray Scanner G2505B (Agilent Technologies, Palo Alto, CA, USA). Log2 ratios were calculated with moving average and are displayed against their position on the genome. Smoothed values were calculated using the January 2005 version of array CGH smooth [13].

RESULTS

Case 1

Tumor localizations and CT scans of patient 1 are shown in Figure 1A and 1B. In March 2005, two lesions were observed in the left lung (2 and 3). Lesion 2 was in retrospect already visible at diagnosis in 2003, and did not respond to chemotherapy. EGFR and KRAS mutation analyses were performed on tumor samples from sites 1, 2 and 3, as well as on tumor cells from lymph nodes number 5 and 10. All tumor sites analyzed contained the same heterozygous, in frame deletion of 15 base pairs in exon 19 of the EGFR gene, leading to the deletion of amino acids E746 to A750. This deletion affects the ATP binding pocket and is related to favorable response to treatment with EGFR TKIs [14,15]. All sites analyzed were wild type for KRAS. Although pathological examination suggested metastatic adenocarcinoma, radiologically the tumors in the left lung were compatible with being second primaries. To get a better insight in the relation of the various lesions, array CGH was performed. Because of the small number of tumor cells left in site 1, this analysis was performed on tumor sites 2 and 3 only. Profiles of the genetic aberrations are depicted in Figure 1C. Log2 tumor to normal DNA copy number ratios are plotted in genomic order. The aberrations are spread over the whole genome, and follow a nearly identical pattern of gains and losses in both sites, indicating that these tumors share the same clonal origin.

Case 2

In Figure 2A and 2B, a schematic diagram of the localization and CT scans of the tumors are presented. The clinical presentation of this patient suggested the presence of two independent lung lesions. However, pathological and immunohistochemical examination could not distinguish
FIGURE 3. Localization and molecular analysis of tumors from patient 3. (A) Localization and time of diagnosis of the various malignant lung lesions of patient 3. (B) CT scans of the thorax in which lesions 1, 2, 3, and 4 can be observed. (C) Chromosomal aberration patterns of site 1, 2 and 4. Tumor DNA of patient 3 was hybridized with normal male reference DNA on a 30K human oligonucleotide array. Log2 ratios are displayed against their position on the genome. Odd and even chromosomes are indicated in light grey and black, respectively. Smoothed values are indicated in dark grey. The chromosomal aberrations of tumor 1 and 2 follow a similar pattern, suggesting they share the same clonal origin. The pattern of chromosomal aberrations of tumor site 4 differs at various sites compared to tumor site 1 and 2, most obviously at chromosome 6 and 7 (arrows). (D) In site 1, the log2 ratios of chromosome 6p, 6q and 7p are +0.08, -0.22 and +0.15, respectively. In site 2, the log2 ratios of chromosome 6p, 6q and 7p are +0.13, -0.18 and +0.12, respectively. Site 4 however, shows a log2 ratio of -0.01 for chromosome 6p-q and the log2 ratio for chromosome 7p is zero. Considering the fact that tumor site 4 harbors a much less altered chromosomal aberration pattern, it is likely that tumor 4 originated from a different clone than tumors 1 and 2.
FIGURE 3. (Continued)
FIGURE 3. (Continued)
between second primary or metastasis. Mutation analysis was performed on tumor material of the thyroid (site 1), the right upper lobe (site 2) and of the left lower lobe (site 3). No mutations in the EGFR gene were found in any of the sites analyzed. Tumor cells of sites 1 and 2 were wild type for KRAS. In contrast, a point mutation in codon 12 of KRAS leading to an amino acid change from glycine to valine (G12V) was detected in the tumor cells of site 3.

Array CGH analysis was performed on tumor material from the two lung lesions (sites 2 and 3). As shown in Figure 2C, the array CGH profiles of these two lesions differ for almost each and every of the 22 chromosomes. In site 3, losses on chromosomes 3, 6, 9, 10, 14 and an amplification on chromosomes 14p can be observed, whereas in tumor 2 these aberrations are not present.

**Case 3**

A schematic representation of the tumor sites and CT scans of patient 3 are shown in Figure 3A and 3B. Pathological examination revealed papillary adenocarcinoma with the same morphology in all tumor sites, which suggests that this is most probably a metastasizing tumor. Mutational analysis of EGFR and KRAS was performed on tumor cells from site 1, 2 and 4. All three lesions were wild type for EGFR. Sites 1 and 2 were also wild type for KRAS, whereas a point mutation (G12V) in this gene was detected in site 4. Since KRAS mutations may be an early event in tumorigenesis [16,17], the KRAS mutation present only in site 4, suggests that site 4 could be a new primary tumor. Array CGH was performed on tumor sites 1, 2 and 4. In Figure 3C, as can be observed from smoothed patterns, it is shown that very few chromosomal aberrations are present in site 4, while in sites 1 and 2 similar patterns of gains and losses can be observed. In particular, gains of 6p and 7p and loss of 6q present in sites 1 and 2 are absent in tumor site 4 (Figure 3 D).

**DISCUSSION**

In order to identify the optimal treatment in patients with multiple lung tumors, the correct diagnosis of primary vs secondary tumors is essential. Routine clinical and pathological evaluation to determine the relationship between the different lesions often is not fully conclusive and a detailed genetic analysis of tumor samples (e.g. for detection of gene mutations, amplifications or deletions) may provide important additional information.

We describe here three NSCLC patients, presenting with multiple lesions and the results of genetic analysis of their tumor samples. The analysis included direct sequencing to detect EGFR and KRAS gene mutations, and array CGH to obtain tumor specific DNA copy number signatures. Previous groups have established that CGH is a valuable tool in determining clonal relations between primary tumors and their metastases of the breast [18], bladder [19], colon [20], and
CHAPTER 4

lung [21]. Other studies in determining relations between multiple tumors in one patient used loss of heterozygosity (LOH) based techniques [22]. In this technique only limited regions of the genome are compared to identify a clonal relation. We applied array-based CGH, capable of a genome wide screening for copy number aberrations, with high resolution and high sensitivity [23].

In patient 1, the same EGFR mutation was found in all lesions. In addition, nearly identical patterns of gains and losses could be observed in lesion 2 and 3. The results of the mutation analysis and the array CGH are consistent with tumors 2 and 3 being metastatic lesions arising from tumor 1. This would also support the view that the occurrence of EGFR mutations are an early event in NSCLC tumorigenesis [24]. In patient 2, clinical diagnosis suggested the presence of two independent tumors. This hypothesis is strengthened by the results of the molecular analysis. The dissimilar chromosomal aberration pattern and the different KRAS mutational status suggest that lesions 2 and 3 are most probably not clonally related, and may have genetically different underlying pathways of tumor development. For patient 3, the results of the array CGH analysis suggests that lesions 1 and 2, but not 4 are clonally related. Thus, whereas lesion 2 (2002) could be a metastasis from lesion 1 (1998), it is most likely that lesion 4 (2006) arose independently. As described in patient 2, the possibility of a different tumorigenic pathway driving the development of tumor 4 is further supported by the fact that a KRAS mutation was only detected in this lesion.

In summary, we show that by using mutation analysis of EGFR and KRAS combined with array CGH, we were able to identify clonal relations between multiple tumors in one patient. Our findings suggest that, in patients presenting with multiple lung lesions, various underlying molecular mechanisms can be responsible for tumor development. The predictive value of specific mutations in EGFR and KRAS genes has already brought us one step closer to more individualized treatment approaches. The cases described here, show that a detailed molecular analysis consisting of both EGFR and KRAS mutation analysis together with array CGH can help clinicians to determine the appropriate stage and treatment for patients with lung cancer presenting with multiple pulmonary nodules.
GENETIC HETEROGENEITY IN MULTIPLE NSCLC LESIONS

REFERENCES


CHAPTER 4


EGFR AND K-RAS MUTATION ANALYSIS IN NON-SMALL CELL LUNG CANCER: COMPARISON OF PARAFFIN EMBEDDED VERSUS FROZEN SPECIMENS

Mariëlle I Gallegos Ruiz, Karijn Floor, Frank Rijmen, Katrien Grünberg, José A Rodriguez and Giuseppe Giaccone

CELLULAR ONCOLOGY 2007
29: 257-264
ABSTRACT

**Background:** Mutational analysis of the Epidermal Growth Factor Receptor (EGFR) and K-ras genes to select non-small cell lung cancer (NSCLC) patients for treatment with novel EGFR tyrosine kinase inhibitors is an appealing possibility currently under investigation. Although frozen tumor tissue would probably be the optimal source for analysis, the most common source of tumor material is fixed and paraffin embedded (FPE) archival specimens. Here, we evaluate how different procedures of tissue sample processing and preservation may affect the outcome of EGFR and K-ras mutation analysis. Furthermore, we compare the sensitivity of the analysis using genomic DNA (gDNA) versus RNA.

**Methods:** We used PCR amplification and direct sequencing to analyze EGFR and K-ras genes in paired FPE and frozen tumor samples corresponding to 47 NSCLC patients. In frozen samples, the analysis was carried out using both gDNA and RNA extracted in parallel.

**Results:** Whereas 100% of frozen samples were successfully amplified, the rate of successful PCR amplification in FPE samples was approximately 50%. We detected three previously described EGFR point mutations in 2 samples. In ten other samples, a K-ras mutation was observed. These mutations were detected in DNA extracted from frozen samples as well as in DNA obtained from FPE tissue. In addition, 10 nucleotide changes, were detected in FPE samples that were not detected in the frozen specimens. Upon re-analysis, these nucleotide changes could not be confirmed and were most likely the result of paraffin embedding and fixation procedures. All mutations found in gDNA were also detected in the corresponding RNA and, in two cases, the presence of the mutant allele was easier to identify by using RNA.

**Conclusions:** Our results indicate that RNA extracted from frozen tissue is the preferred source for EGFR and K-ras mutation testing. When analyzing FPE samples, reducing the size of the amplified fragments would increase PCR success rate, and care should be taken to control for false-positive results.
INTRODUCTION

The Epidermal Growth Factor Receptor (EGFR) plays an important role in a variety of tumor types, including non-small cell lung cancer (NSCLC) [1]. Inhibitors of EGFR tyrosine kinase activity, such as gefitinib and erlotinib, have definite activity in a subset of NSCLC patients [2]. Response rates vary between 10 to 20 percent when used as second- or third line treatment for advanced disease [3,4]. Higher responses to EGFR tyrosine kinase inhibitors (TKIs) have been seen in east-Asian patients, in women, in non-smokers, and in patients with adenocarcinoma [5,6]. Tumors of the majority of the responsive patients possess somatic EGFR mutations in the kinase domain of the receptor [7-9]. These mutations are most commonly short deletions in exon 19 or the L858R point mutation in exon 21, and are present in 10% of Caucasian patients and 30% of NSCLC patients with east-Asian ethnicity [10]. Besides these predominant mutations, some less frequent mutations have been described, i.e. Glu709Gly, Gly719Ser, Ser768Ile, and Leu861Gln [11,12]. In vitro, these "rare" mutations were also shown to be related to higher sensitivity to gefitinib [13]. In contrast to EGFR gene defects, mutations in K-ras predicted primary resistance to EGFR TKIs [14]. Therefore, the use of EGFR and K-ras mutation analysis may potentially have an impact in the selection of patients to be treated with EGFR TKIs.

Although several approaches for analyzing EGFR and K-ras mutations have been described [15-18], the most commonly used approach is direct sequencing of PCR-amplified DNA from tumor specimens. Several studies have used frozen samples as a source of DNA and or RNA for mutational analysis of EGFR [19,20]. These studies used tumor tissue resected from patients with early stage NSCLC. However, in clinical practice, and particularly in patients with advanced disease, the main source of tumor samples is fixed and paraffin embedded (FPE) archival tissue blocks. Fixation procedures can result in DNA damage making these samples difficult to analyze and potentially affecting the results of the analysis [21].

In order to evaluate how different methods of tissue processing and preservation may influence the outcome of EGFR and K-ras mutation analysis, we used PCR amplification and direct DNA sequencing to analyze the mutational status of EGFR (exons 18-21) and K-ras (exons 1 and 2) in a series of paired FPE and frozen tumor samples corresponding to 47 NSCLC patients. In addition, we compared the influence on PCR success rate of sublimate formalin versus neutral buffered formalin in paraffin embedded specimens. Sublimate formalin is a mercury-containing fixative that has been used in some Pathology Departments to obtain a better morphology for immunohistochemistry purposes, but is nowadays commonly replaced by 4% neutral buffered formalin. In the frozen samples, we also compared the sensitivity of the analysis using genomic DNA (gDNA) versus RNA.
MATERIALS AND METHODS

Samples

Resected tumor samples from 47 NSCLC patients of which both frozen and FPE material was available were selected. Of these resected specimens, one piece of the tumor was snap-frozen and stored in liquid nitrogen and another piece was processed for fixation and paraffin embedding. 20 of the FPE tissues had been fixed in 4% neutral buffered formalin and the remaining 27 had been fixed using sublimate formalin, a mercury-containing fixative. The collection date of the samples ranged from 1992 to 2002. The range for neutral buffered formalin samples is 1992-2002 and the range of the sublimate formalin is from 1994-2000.

DNA and RNA isolation

For FPE samples, a 4 μm-thick section was stained with hematoxylin-eosin and examined by a pathologist (KG) to select areas of tissue containing at least 50% tumor cells. These areas were macro-dissected from 15 unstained consecutive sections of 10 μm, and DNA isolation was performed using the QIAamp DNA mini kit as indicated by the manufacturer (Qiagen, Venlo, The Netherlands). For frozen samples, cryo-sections were verified by the pathologist to contain at least 50% of tumor cells. Genomic DNA and RNA was extracted from each sample using Trizol (Life Technologies, Breda, The Netherlands). Tissue samples were homogenized in Trizol (1 ml per 50-100 mg of tissue), and chloroform (200 μl per ml of Trizol) was added. After incubation at room temperature for 3 minutes and centrifugation for 10 minutes at 12000 g at 4°C, the aqueous layer containing the RNA was transferred to a new vial and the interphase was stored for subsequent DNA isolation. RNA was precipitated by addition of isopropanol and centrifugation. The pellet was washed with 1 ml 70% ethanol, air-dried, and dissolved in RNase-free water.

Table 1. Success rate of PCR on DNA derived from frozen and sublimate formalin or neutral buffered formalin fixed paraffin embedded NSCLC specimens

<table>
<thead>
<tr>
<th></th>
<th>Frozen</th>
<th>Paraffin</th>
<th>SF</th>
<th>NBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR Exon 18 (248 bp)</td>
<td>47 (100)</td>
<td>20 (43)</td>
<td>6  (30)</td>
<td>14 (70)</td>
</tr>
<tr>
<td>EGFR Exon 19 (225 bp)</td>
<td>47 (100)</td>
<td>34 (72)</td>
<td>18 (53)</td>
<td>16 (47)</td>
</tr>
<tr>
<td>EGFR Exon 20 (325 bp)</td>
<td>47 (100)</td>
<td>21 (45)</td>
<td>7  (23)</td>
<td>14 (47)</td>
</tr>
<tr>
<td>EGFR Exon 21 (301 bp)</td>
<td>47 (100)</td>
<td>22 (42)</td>
<td>7  (23)</td>
<td>15 (46)</td>
</tr>
<tr>
<td>Kras Exon 1 (253 bp)</td>
<td>47 (100)</td>
<td>19 (40)</td>
<td>5  (26)</td>
<td>14 (74)</td>
</tr>
<tr>
<td>Kras Exon 2 (295 bp)</td>
<td>47 (100)</td>
<td>9  (19)</td>
<td>0  (0)</td>
<td>9  (100)</td>
</tr>
</tbody>
</table>

Frozen, number of samples on which PCR was successful in frozen specimens; Paraffin, number of samples on which PCR was successful in paraffin specimens; SF, samples fixed in Sublimate Formalin; NBF, samples fixed in Neutral Buffered Formalin.
For DNA isolation, 100% ethanol (0.3 ml per ml of Trizol) was added to the stored Trizol interphase. After mixing and incubation at room temperature for 3 minutes, samples were centrifuged for 10 minutes at 20000 g at 4°C. The DNA pellet was washed twice in 1 ml 10% ethanol containing 0.1 M sodium citrate and once with 70% ethanol. Finally, the pellet was air-dried and dissolved in milliQ water.

**PCR amplification and DNA sequencing**

We used 100 ng of genomic DNA, derived from tumor cells, as template in nested PCR reactions to amplify DNA fragments corresponding to exons 18-21 of EGFR and exons 1 and 2 of K-ras. The PCR protocol and the sets of primers have been described in detail previously [22]. Additionally, total RNA extracted from frozen samples was used as template to amplify a single PCR fragment encompassing EGFR exons 18-21 using the primers designed by Kosaka et al. [23], for 5'-GGTGCGGGAGAGAGGCCTGCT-3' and REV 5'-ATCTTCAGAGTCCTTAACTC-3'. To this end, the Qiagen one-step reverse transcription PCR kit was used according to the manufacturer's instructions. PCR products were purified using a presequencing kit (Amersham Biosciences, Roosendaal, The Netherlands) and sequenced with both forward and reverse primers using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA), with the ABI PRISM™ 3100 Genetic analyzer (Applied Biosystems). Mutations were confirmed by sequencing independent PCR products, on DNA derived from the tumor cells.

**Statistical Analysis**

Logistic regression techniques were used to analyse the data. The generalized estimating equations (GEE) method [24] was applied to account for the variations in success rate of various DNA fragments derived from one sample.

![Sequence chromatograms for sample #3977909 showing a wild type sequence when DNA from frozen tissue was used as template for PCR, whereas when DNA extracted from the fixed and paraffin embedded (FPE) sample was used as template for PCR, a mutation in exon 20 could be observed, Gly779Ser (arrow). Re-analyzing this fixed and paraffin embedded sample revealed a wild type sequence. A colored version of this figure can be found in appendix C – page 197.](image)
RESULTS

Rate of successful PCR amplification using genomic DNA extracted from FPE versus frozen samples

Frozen and FPE tissue samples corresponding to 47 NSCLC patients were selected based on availability. Samples were collected between 1992 and 2002. Using genomic DNA extracted from either the FPE or the frozen tissue as template, nested PCR reactions were carried out to amplify 6 separate DNA fragments encompassing EGFR exons 18-21 and K-ras exons 1 and 2. DNA extracted from frozen samples was successfully amplified in 100% of all 47 samples for each exon. The success rate of PCR amplification for DNA extracted from FPE samples however, was much lower. Amplification of EGFR exon 18 (248 bp) was successful in 43%, in exon 19 (225 bp) the percentage of successful amplifications was 72%, PCR for EGFR exon 20 (325 bp) was successful in 45% and PCR products for EGFR exon 21 (301 bp) were formed in 42% of cases. For Kras, PCR for exon 1 (253 bp) was successful in 40% and exon 2 (295 bp) in only 19% of cases (Table 1). Thus, rate of successful PCR amplifications appeared to be, at least partially, related to the size of the product, which is as expected. These results suggest that the use of shorter amplification products may partially circumvent a low PCR success rate when working with FPE samples, which likely contain fragmented DNA. To evaluate this possibility, we designed a new set of primers to reanalyze K-ras exon 2. Indeed, by reducing the size of the amplified DNA fragment from 295 bp to 235 bp the success rate was increased from 19% to 61%.

Effect of different fixatives on the rate of successful PCR amplifications using DNA extracted from FPE samples

27 of the 47 FPE tissue specimens used in this study had been fixed in sublimate formalin and 20 in 4% neutral buffered formalin. Success rates of PCR in sublimate formalin (SF) fixated samples versus samples fixed in neutral buffered formalin (NBF) are indicated in Table 1. To investigate the potential effect of fixative on the PCR success rate, a logistic regression analysis was performed. The generalized estimation equations method with exchangeable working correlation structure was used to take into account the dependencies between fragments amplified from the same sample. The analysis revealed that the success rate was significantly higher in samples fixed with neutral buffered formalin than in samples fixed with sublimate formalin (75% vs. 33%, P=.001). We furthermore evaluated the influence of storage time on rate of successful amplifications but no significant effect was observed, thereby also excluding any possible confounding between storage time and type of fixative used.
Detection of mutations and single nucleotide polymorphisms (SNPs) in DNA from FPE versus frozen samples

Three previously described EGFR point mutations \[25\], Pro848Leu, Ser768Ile and Leu861Gln, were detected in two frozen samples (one sample contained both Ser768Ile and Leu861Gln). No results were obtained in the matching FPE samples because of PCR failure. The EGFR single nucleotide polymorphism (SNP) Arg836Arg in exon 21 was found in 4 cases. In two of these patients, PCR was successful in the corresponding FPE samples and the Arg836Arg SNP was detected. A second SNP (Gln787Gln) in exon 20 was detected in 42 out of 47 patients using frozen samples. The SNP was also detected in the corresponding FPE samples in those twenty cases for which a successful PCR amplification of exon 20 was possible. We identified 10 nucleotide changes initially thought to be mutations in FPE samples that were not detected in the corresponding frozen sample (Figure 1). Upon re-analysis of these samples, a wild type sequence was observed, indicating that these nucleotide changes were most probably artifacts. As detailed in Table 2, all artifactual mutations resulted from C>T or G>A transitions. A total of 10 artifactual nucleotide changes were observed in 7 samples. 5 samples harboring the artifactual changes were fixed in sublimate formalin and each contained 1 nucleotide change. Two samples were fixed in neutral buffered formalin and contained 2 and 3 nucleotide changes per sample, respectively. Therefore, the type of fixative does not appear to have a clear influence on presence of artifactual mutations, although the possibility of drawing statistically meaningful conclusions is hampered by the limited amount of samples. On the other hand, 7 codon 12 and 3 codon 13 K-ras mutations were detected using DNA extracted from frozen samples. These
mutations were also detected in the corresponding FPE samples in the 4 cases for which a successful PCR amplification of K-ras exon 1 was possible. No artifactual mutations in K-ras were observed in DNA extracted from FPE samples.

Sensitivity of EGFR analysis using genomic DNA versus RNA

To determine whether the use of genomic DNA or RNA may influence the sensitivity of the analysis, we compared the sequencing results obtained using either type of template. Genomic DNA and RNA were extracted in parallel from 30 frozen tissue samples using Trizol as described in the Materials and Methods section. In every case, the EGFR point mutations (Pro848Leu, Ser768Ile and Leu861Gln) and SNPs (Arg836Arg and Gln787Gln) detected using genomic DNA as template were also detected using RNA. The chromatograms obtained using either template were similar in the case of the Pro848Leu mutation and the SNPs (see for SNP Gln787Gln, Figure 2a). However, the peaks corresponding to the mutant allele for Ser768Ile and Leu861Gln were more evident in the sequencing chromatogram from RNA (Figure 2b and c). This observation suggests that, at least in some cases, RNA would be a better template for EGFR mutation analysis.

![Figure 2](image.png)

**FIGURE 2.** (A) Sequencing EGFR from gDNA and RNA in sample #9446712 showing the single nucleotide polymorphism Gln787Gln and (B-C) the EGFR point mutations Ser768Ile and Leu861Gln. Whereas the wildtype to mutant allele ratio in the polymorphism is comparable in the chromatograms from gDNA versus RNA (A - arrows), both EGFR point mutation peaks are elevated in the chromatograms from RNA compared to the chromatograms from gDNA (B and C - arrows). A colored version of this figure can be found in appendix C – page 197.
CHAPTER 5

DISCUSSION

This study was conducted to compare the sensitivity and specificity of detecting EGFR and K-ras mutations in RNA and DNA from frozen and DNA from fixed and paraffin embedded (FPE) and frozen lung cancer specimens, using direct sequencing. In addition, two different types of fixatives were compared for the effect on success rate of PCR.

100% of frozen samples could be successfully analyzed whereas only ~50% of PCRs on FPE samples succeeded, which was partly correlated with PCR product size. Previous reports identifying EGFR mutations have also used paraffin embedded diagnostic specimens and could successfully amplify PCR products with a size of approximately 400 bp, although the longest PCR product we obtained was 325 bp [26,27]. We show that type of fixative was a main reason for PCR failure. In neutral buffered formalin fixed samples, the rate of successful PCR amplifications was 75% whereas only 33% of samples fixed in sublimate formalin were successfully amplified.

Another factor that could have influenced the PCR success is the time of fixation. As shown by Inoue et al [28], fixation for 2-3 days resulted in a rate of PCR failure of 56%, whereas fixation for 1 day does not appear to damage the DNA. We were unable to determine if this has been of influence in our study, since no standard fixation time is applied in our Pathology department. Although one might also expect the time of storage could influence PCR success rate, this was not of influence in our sample set.

The EGFR point mutation Pro848Leu and SNPs Arg836Arg and Gln787Gln could be detected with the same accuracy in DNA and RNA. However, two EGFR point mutations, Ser768Ile and Leu861Gln were detected with greater accuracy in RNA than in DNA. In particular the Leu861Gln mutation was almost undetectable in DNA whereas the mutant:wildtype allele distribution was ~1:1 in RNA. Since the Gln787Gln SNP in this sample showed comparable distribution of mutant and wild type alleles in RNA and DNA, this is most probably the result of elevated expression of the EGFR mutant alleles. Indeed it has been described by others that the EGFR mutant allele can be selectively amplified [29]. Thus, RNA seems to be a better source for sequencing than DNA for the detection of certain EGFR mutations. In contrast to EGFR mutations, K-ras mutations were detected with the same accuracy using DNA and RNA. This observation may stem for the fact that whereas amplification of mutated EGFR alleles has been previously described [30], mutated K-ras alleles are not usually amplified. Tsao et al reported the presence of 24 novel mutations in 110 patients from the BR21 study [31]. Since these mutations were not confirmed by repeated analysis, the mutations found could have been the result of mutational artifacts due to fixation and paraffin embedding. In our study, 10 novel nucleotide changes were found in FPE samples that were not observed in matching frozen samples. Upon re-amplification and repeated sequencing, these 10 novel nucleotide changes could not be confirmed. As discussed by
Marchetti et al [32], fixation and embedding can result in deamination of cytosine and adenine, leading to the generation of uracil or hypoxanthine, respectively. This may result in the artifactual detection of C>T or G>A transitions, which were exactly the types of non-reproducible changes we detected in some FPE samples. In case these nucleotide changes are indeed the result of fixation, one would expect to detect an equal amount of such artifacts on other locations upon re-analysis. We re-analyzed the samples bearing the initial nucleotide changes and did not observe any additional artifacts on other locations, which is most likely due to the smaller chance on such changes in a smaller sample set. Since these nucleotide changes were only observed in the paraffin samples and could not be confirmed by an independent analysis we assume these nucleotide changes have been the result of paraffin embedding and fixation of the samples. The number of amplifiable templates decreases with increasing product size, making these templates more vulnerable to artifactual amplifications. Of 10 artifactual nucleotide changes that were observed, indeed 5 (50%) were found in the longest product (exon 20, 325 bp). However, 2 artifactual nucleotide changes were observed in the shortest product (exon 19, 225 bp), 2 in exon 21 (301 bp) and 1 in exon 18 (248 bp). A remarkable finding is the fact that these nucleotide changes were observed only in the EGFR templates and not in K-ras. Although the total length of DNA sequence analyzed is longer for EGFR (564 bp) than for K-Ras (290 bp), this is unlikely to wholly account for the difference in artifacts. Thus, there appear to be other factors, besides the size of the PCR product, which may be related to the detection of sequencing artefacts.

In conclusion, this study highlights the technical challenges in detecting EGFR and K-ras mutations and the need for a standardized fixation protocol to allow molecular studies to be completed successfully. When working with DNA derived from FPE, primers should be designed to amplify DNA fragments as short as possible and each nucleotide change should be confirmed using an independent PCR-amplified product. For some EGFR mutations, RNA seems to be a superior source for direct sequencing than gDNA.
REFERENCES


EGFR AND KRAS MUTATION ANALYSIS IN NON-SMALL CELL LUNG CANCER


ACKNOWLEDGEMENTS

Reprinted from Cellular Oncology Volume 29, Number 3, Pages 257-264, (2007), with permission from IOS Press.
CHAPTER 6

EGFR GENE COPY NUMBER DETECTION IN NON-SMALL CELL LUNG CANCER; A COMPARISON OF FLUORESCENT IN SITU HYBRIDIZATION AND CHROMOGENIC IN SITU HYBRIDIZATION

Mariëlle I Gallegos Ruiz, Karijn Floor, Wim Vos, Katrien Grünberg, Gerrit A Meijer, José A Rodriguez and Giuseppe Giaccone

HISTOPATHOLOGY 2007
51: 631-637
ABSTRACT

Aims: The Epidermal Growth Factor Receptor (EGFR) is an important target for anti-cancer therapy. In non-small cell lung cancer (NSCLC), mutations in the tyrosine kinase domain of EGFR and EGFR gene copy number have been demonstrated to identify patients who most likely benefit from EGFR tyrosine kinase inhibitors. EGFR gene copy number has been mainly assessed by Fluorescent In Situ Hybridization (FISH), a method requiring the use of a fluorescence microscope and may be hampered by the rapid fading of the fluorescent signal. These limitations of FISH can be overcome by using Chromogenic In Situ Hybridization (CISH). To test the applicability of CISH for EGFR gene copy number testing in NSCLC, we performed a comparison of CISH and FISH.

Methods and Results: A total of 58 formalin-fixed and frozen NSCLC tissue samples were collected on which both CISH and FISH were performed. A high concordance was found in the assessment of EGFR copy number between observers and between techniques (Kappa coefficient = 0.64-0.76). CISH seems the ideal technique for paraffin sections whereas FISH is favourable for frozen material.

Conclusions: We show that CISH can be a suitable alternative strategy for FISH in determining EGFR gene copy number in NSCLC patients.
INTRODUCTION

The Epidermal Growth Factor Receptor (EGFR) is a proto-oncogene located on chromosome band 7p12. EGFR is overexpressed in several types of cancers including non-small cell lung cancer (NSCLC). Various therapeutic approaches targeting EGFR have been approved by the US Food and Drug Administration. For example, cetuximab, a monoclonal antibody against EGFR is registered for treatment of patients with advanced colorectal cancer, and gefitinib and erlotinib are tyrosine kinase inhibitors (TKI) of EGFR registered for use in patients with locally advanced or metastatic NSCLC who failed chemotherapy. Both gefitinib and erlotinib have been shown to be effective in a small subset of NSCLC patients. Much effort is currently ongoing in identifying those patients who will most likely benefit from anti-EGFR TKI treatment. In this respect, somatic mutations in the tyrosine kinase domain of EGFR have been shown to correlate with clinical benefit [1-3]. Although EGFR mutations predict response to TKI treatment in NSCLC patients, impact on survival is still unclear [4-7] and awaits further confirmation.

Various groups have reported that EGFR copy number, as assessed by Fluorescent In Situ Hybridisation (FISH), is a better prognostic marker than the presence of EGFR mutations, for NSCLC patients treated with TKIs. Cappuzzo et al. [8] showed that high polysomy and amplification of the EGFR gene predicted a favourable response to gefitinib and were correlated with prolonged survival. In the BR21 study, high EGFR copy number was associated with longer survival in patients treated with erlotinib [9]. Furthermore, Hirsch et al found a higher response rate in patients treated with gefitinib in FISH positive patients and high EGFR copy number was associated to improved survival compared to FISH negative patients [6]. In contrast to these reports using FISH, other studies using quantitative PCR failed to obtain predictive information from determining EGFR copy number [10,11]. In summary, these data indicate that EGFR gene copy number assessed by FISH is an important predictive and prognostic biomarker for NSCLC patients treated with EGFR TKIs.

Disadvantages of determining EGFR gene copy number by FISH is that the fluorescent signal is not stable, the signal may fade away in a few weeks and morphology can be hard to detect. In addition, the need for a fluorescence microscope makes it impractical for routine use. An additional problem of FISH is that there are currently no standardized interpretation systems so interchanging results from various laboratories will not always be reliable. Some of these drawbacks can be overcome by Chromogenic In Situ Hybridization (CISH). CISH is a modified non-fluorescence-based in situ hybridization technique in which the signal is not visualized by using a fluorescently labelled probe but using an immunoperoxidase reaction [12]. Therefore, gene copy numbers can be determined by conventional bright field microscopy. Additionally, the CISH signal does not fade, and this allows pathologists to include the scoring of CISH in their
EGFR GENE COPY NUMBER DETECTION IN NON-SMALL CELL LUNG CANCER

routine work. The cells are counterstained with haematoxylin, which allows a better distinction of tumour cells from non-tumour stroma and other cells, compared to the counterstaining used in FISH. Moreover, most laboratories are more familiar with chromogenic signals and bright field microscopy than with fluorescent microscopy, which could allow a better comparison of results between different laboratories.

Implementing CISH for the detection of EGFR copy number may be of practical value in the management of NSCLC patients treated with EGFR inhibitors. Since to date, most studies have been performed with FISH, we performed a comparison of FISH and CISH for the detection of EGFR gene copy number in 58 NSCLC specimens.

MATERIALS AND METHODS

Samples
Formalin-fixed, paraffin-embedded (FFPE) diagnostic specimens of NSCLC patients and frozen tumour samples from resected NSCLC patients were used for this study. The use of these tissues was in accordance with local ethical guidelines. Haematoxylin and eosin stained slides were revised by a pathologist (KG). A total of 58 tumours on which CISH was successfully performed were collected, 31 FFPE samples (15 resection specimens, 8 primary tumour biopsies, 3 lymph node biopsies, 3 brain metastases, 1 liver metastasis, 1 pleural fluid) and 27 frozen samples (all resection specimens). Composition of histological subtypes is; 43% adenocarcinomas, 39% squamous cell carcinomas, 9% large cell carcinomas and 9% others.

Chromogenic In Situ Hybridization (CISH)
4 μm frozen sections were fixed with methanol/acetic acid (3:1). As pre-treatment, frozen tissues were digested for 2 min with 0.01% pepsin/0.2N HCl at 37°C, and incubated for 10 min in 50mM MgCl2/PBS followed by 10 min in 50mM MgCl2/3,7% formaldehyde/PBS. After 2h incubation with 70% formamide/0.6x SSC the sections were dehydrated with alcohol. 4 μm sections cut from FFPE tissue blocks were incubated at 64°C for 2-4h, deparaffinized with xylene and re-hydrated with alcohol. As a pretreatment the FFPE tissues were boiled in 1mM EDTA/Tris pH9 using a microwave (5 min at 1200 W followed by 10 min at 360 W). After cooling down, the sections were digested for 10 min with 0.01% pepsin/0.2N HCl at 37°C, and dehydrated with alcohol. After the different pretreatments described above, frozen and paraffin sections were processed in the same manner. 10-15 μl Spot-Light EGFR probe (Zymed, Sanbio BV, Uden The Netherlands) were applied to each section, the tissue was covered with a coverslip, sealed with rubber cement, and the slides were incubated for 10 min at 80°C for codenaturation of probe and chromosomal DNA. Hybridization was carried out in a humidified chamber for 16-24h at 37°C. After
hybridization, the sections were washed with 0.5x SSC at 75°C for 5 min. Sections were then washed with PBS/0.05% Tween 20. In the case of FFPE tissues, endogenous peroxidase was blocked by incubation with 3% H2O2/PBS for 10 min, followed by a 10 min wash with PBS. After pre-incubation with normal goat serum (1:20), tissues were incubated for 1h with a 1:100 dilution of mouse anti-digoxogenin antibody (clone DI-22, Sigma Aldrich Chemie BV, Zwijndrecht, The Netherlands) and the bound antibody was visualized using the DAKO Envision™ system (Dakocytomation, Heverlee, Belgium). Finally, the tissue sections were counterstained with haematoxylin and covered with Depex (Gurr, Poole, United Kindom).

Table 1. Success rates for FISH and CISH in relation to source of the sample

<table>
<thead>
<tr>
<th>Number of samples defined by success rate</th>
<th>Frozen (n=27)</th>
<th>Paraffin (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FISH</td>
<td>CISH</td>
<td>FISH</td>
</tr>
<tr>
<td>Easily assessable</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>Difficult to assess</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Not assessable</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

FISH, Fluorescent In Situ Hybridization; CISH, Chromogenic In situ Hybridization

Fluorescent In Situ Hybridization (FISH)

Frozen sections were fixed and pretreated as described for CISH. FFPE tissue sections, were incubated at 64°C for 2-4h, deparaffinized with xylene and re-hydrated with alcohol. Next, FFPE sections were incubated with 2x SSC (pH7) at 75°C for 20-25 min, digested for 17 min using 0.25 mg/ml proteinase K in 2x SSC (pH 7) at 37°C, washed 5 min with 2x SSC and dehydrated with alcohol.

After pretreatment, 10-15 LSI EGFR Spectrum Orange/CEP7 Spectrum Green probe (Vysis, Abbott Laboratories, Downers Grove, IL, USA) were applied, and the section was covered with a coverslip and sealed with rubber cement. Following a denaturation step at 80°C for 10 min, slides were placed in a humidified chamber at 37°C for 20-24h. Then, sections were washed with 1.5M urea/0.1x SSC at 45°C for 30 min, and with 2x SSC for 2 min. Finally, sections were counterstained with 4',6-Diamidino-2-phenylindole (DAPI, Sanbio BV) and mounted using Vectashield (Brunschwig Chemie, Amsterdam, The Netherlands).
Scoring

Samples prepared for CISH were evaluated with an Olympus BX50F optic microscope (Olympus Optical Co Ltd, Tokyo, Japan) using a 40x Plan objective (NA=0.65). Samples prepared for FISH were evaluated using a Leica DMRA fluorescent microscope (Leica Microsystems BV, Wetzlar Germany) with a 60x PL Fluotar oil immersion objective (NA=1.40). Scoring was done by two observers (KF and MGR) in a blinded manner on coded slides. For every sample, the complete section was screened for homo/heterogeneity of the FISH and CISH signals. In at least three representative microscopic fields, the signals in 200 tumour cells were counted. The number of cells having 0, 1, 2, 3, 4, 5 >6 signals or clusters was noted and samples were categorized as described previously [6]. Samples were considered positive when • 4 dots per nucleus were present in • 40% of tumour cells, or when tight EGFR gene clusters were present in ≥10% of cells.

Statistical analysis

Agreement between observers and between techniques was determined using the Kappa statistics.

Table 2. Cases with different outcome between observers for CISH

<table>
<thead>
<tr>
<th>Case #</th>
<th>OB1</th>
<th>OB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>30 (-)</td>
<td>43 (+)</td>
</tr>
<tr>
<td>36</td>
<td>23 (-)</td>
<td>44 (+)</td>
</tr>
<tr>
<td>44</td>
<td>25 (-)</td>
<td>66 (+)</td>
</tr>
<tr>
<td>39</td>
<td>43 (+)</td>
<td>13 (-)</td>
</tr>
</tbody>
</table>

CISH, Chromogenic In situ Hybridization; OB1, Observer 1 (KF); OB2, Observer 2 (MGR); (+), CISH positive; (-), CISH negative

Table 3. Correlation between FISH and CISH results for both observers

<table>
<thead>
<tr>
<th>CISH</th>
<th>FISH</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observer 1</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>2</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Observer 2</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>2</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

FISH, Fluorescent In Situ Hybridization; CISH, Chromogenic In situ Hybridization
CHAPTER 6

RESULTS

Success rate
CISH and FISH analyses were performed on 58 NSCLC tissue samples that have been preserved either frozen in liquid nitrogen (n=27) or as FFPE tissue blocks (n=31). As illustrated in Table 1, all the frozen samples initially analyzed by CISH could successfully be scored using FISH. In contrast, 45% (14 out of 31) of the FFPE cases successfully analyzed by CISH could not subsequently be analyzed by FISH. Reasons of failure were auto fluorescence (4 cases), complete detachment of tissue (5 cases), unrecognizable morphology (4 cases) or a low signal (1 case). Although the results of the CISH analysis were interpretable in all the frozen sections, we noted that the assessment was difficult in some samples due to high background. In summary, we found that the percentage of easily assessable samples is higher using CISH vs. FISH in FFPE tissue samples, but FISH is often easier to interpret than CISH in frozen samples.

Concordance between observers
A 93% concordance between the two observers in the scoring of the CISH samples was found (Kappa coefficient = 0.74). In Table 2, the percentages of cells with ≥4 EGFR gene copies are shown for the four discordant cases which were classified as positive by one of the observers and as negative by the other. Taking into account that a sample was considered positive when ≥4 EGFR gene copies are detected in ≥40% of tumour cells as previously described [6], it can be seen that three out of four discordant cases were due to borderline positivity. In one case, however one observer found only 25% of cells with ≥4 copies per cell whereas the other observer found cells with ≥4 copies in 66% of cells, which was most likely due to the heterogeneity of the sample. The concordance between the two observers was 100% (Kappa coefficient = 1) in the scoring of the 44 cases assessable by FISH analysis.

Concordance between CISH and FISH
As shown in Table 3, the concordance between the CISH and FISH results was 93% for observer 1 (Kappa coefficient = 0.64) and 95% for observer 2 (Kappa coefficient = 0.76). Again, the discordant cases were due to the strict cut-off value, scoring negative with one of the techniques and borderline positive with the other. For example, one case, considered FISH negative with 17% of tumour cells having ≥4 EGFR gene copies, was found positive for CISH with 43% of cells with ≥4 copies. Conversely, two cases were CISH negative, while a borderline positive result (45% and 50% of cells with 4 copies respectively) with FISH was found. In figure 1A-F examples of concordant samples as determined by FISH and CISH are shown.
FIGURE 1. (A) Frozen tumor section negative for EGFR gene amplification showing 2-3 copies of the EGFR gene by Chromogenic In Situ Hybridization and (B) by Fluorescent In Situ Hybridization. (C) Frozen tumor section showing high EGFR gene copy number, detected by Chromogenic In Situ Hybridization. (D) A similar result is found with Fluorescent In Situ Hybridisation, showing 6-10 dots per nucleus defined as high polysomy. Surrounding lymphocytes show two copies. (E) Paraffin tumor section positive for EGFR gene amplification showing EGFR gene clusters by Chromogenic In Situ Hybridization and (F) by Fluorescent In Situ Hybridization. A colored version of this figure can be found in appendix C – page 200.
DISCUSSION

In order to evaluate the applicability of CISH for EGFR gene copy number testing in NSCLC, we performed a side-to-side comparison between CISH and FISH in 58 NSCLC samples. CISH has been shown to give similar results to FISH for Her2/neu amplification in breast cancer [13-15]. To our knowledge, a comparison of both techniques for the detection of EGFR in NSCLC has not been performed before. We show a good concordance between the two techniques implying that CISH could be a useful alternative for FISH to detect EGFR gene copy number in NSCLC patients. The incidence of high EGFR gene copy number in NSCLC ranges from 8-10% [16,17]. Several studies have shown correlation between high EGFR gene copy number and response to EGFR TKI treatment [6,18-20]. EGFR gene copy number in NSCLC in relation to treatment with EGFR TKIs has been studied by both Fluorescent In Situ Hybridization and quantitative Real Time PCR (qRT-PCR). No relation could be observed between response to anti-EGFR TKI treatment and copy number as determined by quantitative RT-PCR [11,21], except for one report from Japan [22]. Reasons for this can be that for qRT-PCR the copy number of the gene relative to a reference gene is calculated, whereas no information on the copy number of the reference gene itself is obtained. In addition, in heterogeneous cases, positive signals from amplified cells will be diluted by (non-malignant) cells with a lower level of EGFR gene copy number when using qRT-PCR. Therefore, a detection method using preserved tissue with intact morphology is likely to be more sensitive in assessing gene copy number. Thus, in heterogeneous tumours such as NSCLC, an in situ technique such as FISH, may be the best method to assess EGFR gene copy number. The major disadvantages of FISH include the need of an expensive fluorescent microscope with multi-bandpass filters and the fact that the signal fades within a few weeks. These disadvantages can be overcome by a technique which uses an immunoperoxidase reaction to visualize the signal, Chromogenic In Situ Hybridisation (CISH). With CISH, a bright field microscope can be used and the signal is permanent. In addition, the cells are counterstained with haematoxylin which allows a better distinction of tumour cells from surrounding tissue. A dual-colour CISH approach has been recently introduced [17], for the detection of Her-2/neu. We used a single colour CISH probe to detect EGFR gene copy number, which could be a potential disadvantage of CISH over FISH. In FISH a dual labelled probe is used to visualize the EGFR copy number as compared to the chromosome 7 centromere, thereby allowing to distinguish between aneusomy and true EGFR gene amplification. However, since both balanced polysomy of chromosome 7 and true amplification of EGFR were shown to be correlated with response to TKI and survival in NSCLC patients, this distinction may be less important here.

Only 76% of 58 samples were adequately assessable for FISH, and the reasons for this were auto fluorescence, low signal, unrecognizable morphology or complete detachment of tissue. In
contrast, all samples could be interpreted with CISH, although we selected the samples based on successful hybridizations for CISH. It is known that duration and source of fixation can have influence on CISH success. We showed that for CISH, results were easier to interpret in paraffin than in frozen sections. In contrast, FISH scoring was easier in frozen sections. In comparing the two techniques, several discordant cases were found. We employed the scoring system described by others who showed correlation between FISH+ and FISH- samples and anti-EGFR TKI treatment [6,23]. In this scoring system, the sample is considered positive when ≥40% of cells contain ≥4 copies per tumour core or when tight EGFR gene clusters are present in ≥10% of cells. In most discordant cases the percentage of cells with ≥4 copies per cell was just above 40%. This indicates that additional research has to be performed in determining the appropriate cut-off between high and low polysomy, in defining to what extent the EGFR gene copy number influences response to anti-EGFR TKI treatment. In addition, effort should be made in standardizing the interpretation as much as possible to allow interlaboratory comparison of results [24].

The detection of EGFR copy number in NSCLC patients, could be of importance for diagnostic purposes in the future, to define the subgroup of NSCLC patients most likely to benefit from EGFR TKI treatment. We here show that CISH could be very useful for routine diagnostic practice. A good concordance was found between observers for CISH (Kappa coefficient = 0.74) and for FISH (Kappa coefficient = 1), while the agreement between the two techniques was high (Kappa coefficient = 0.64-0.76). In addition, EGFR copy number detected by CISH has several technical advantages over FISH.

In conclusion, CISH may be a good alternative strategy for FISH in determining EGFR gene copy number in patients with NSCLC, especially in paraffin embedded samples.
REFERENCES


ACKNOWLEDGEMENTS

CHAPTER 7

FUNCTIONAL ANALYSIS OF CANCER-ASSOCIATED EGFR MUTANTS USING A CELLULAR ASSAY WITH YFP-TAGGED EGFR INTRACELLULAR DOMAIN

Matheus M de Gunst, Mariëlle I Gallegos Ruiz, Giuseppe Giaccone and José A Rodriguez

MOLECULAR CANCER 2007
6: 56
ABSTRACT

Background: The presence of EGFR kinase domain mutations in a subset of NSCLC patients correlates with the response to treatment with the EGFR tyrosine kinase inhibitors gefitinib and erlotinib. Although most EGFR mutations detected are short deletions in exon 19 or the L858R point mutation in exon 21, more than 75 different EGFR kinase domain residues have been reported to be altered in NSCLC patients. The phenotypical consequences of different EGFR mutations may vary dramatically, but the majority of uncommon EGFR mutations have never been functionally evaluated.

Results: We demonstrate that the relative kinase activity and erlotinib sensitivity of different EGFR mutants can be readily evaluated using transfection of an YFP-tagged fragment of the EGFR intracellular domain (YFP-EGFR-ICD), followed by immunofluorescence microscopy analysis. Using this assay, we show that the exon 20 insertions Ins770SVD and Ins774HV confer increased kinase activity, but no erlotinib sensitivity. We also show that, in contrast to the common L858R mutation, the uncommon exon 21 point mutations P848L and A859T appear to behave like functionally silent polymorphisms.

Conclusion: The ability to rapidly obtain functional information on EGFR variants of unknown relevance using the YFP-EGFR-ICD assay might prove important in the future for the management of NSCLC patients bearing uncommon EGFR mutations. In addition, our assay may be used to determine the response of resistant EGFR mutants to novel second-generation TKIs.
BACKGROUND

Approximately 80% of lung cancers, the most frequently diagnosed type of human tumor, are classified as non-small cell lung cancer (NSCLC). Novel therapeutic agents for the treatment of NSCLC patients are currently under intense experimental and clinical investigation, with the goal of increasing their antitumor effect while reducing general toxicity. These agents specifically target cellular pathways necessary for the survival of cancer cells. The epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase (TK) whose activation initiates signal transduction through critical cellular pathways, such as those mediated by AKT and ERK, and thus plays an important role in controlling cell homeostasis [1]. EGFR is overexpressed or aberrantly activated in different types of human tumors, contributing to the malignant phenotype of cancer cells, and targeted inactivation of EGFR is being intensively explored as a cancer therapeutic approach [2]. As a result of these investigations, several small-molecule EGFR tyrosine-kinase inhibitors (TKIs), such as gefitinib and erlotinib, have been developed and are currently available in the clinic. In large clinical studies of gefitinib and erlotinib, it became apparent that a minor subset of NSCLC patients is extremely sensitive to treatment with EGFR-TKIs (reviewed in [3]). Subsequently, the analysis of EGFR gene sequence revealed the presence of somatic mutations in the kinase domain of the receptor in most responding patients [4-6]. The association between the presence of EGFR mutations and response to TKIs has been confirmed through the analysis of thousands of NSCLC tumor samples worldwide. These results raise the possibility that EGFR mutational analysis may be implemented for the management of NSCLC patients [7].

Approximately 80% of the EGFR mutations detected are short deletions in exon 19 affecting the amino acid sequence ELREA (Del746-750), or a point mutation in exon 21 resulting in the amino acid change L858R. However, the data accumulated in the past three years have uncovered the large allelic heterogeneity that characterizes EGFR kinase mutations. Thus, a survey of the COSMIC mutation database (http://www.sanger.ac.uk/genetics/CGP/cosmic/) shows that more than 75 different EGFR kinase domain residues have been reported to be altered in NSCLC patients.

The functional characteristics of the two most common types of EGFR alterations, the exon 19 deletions and the L858R point mutation, have been studied in detail using biochemical assays, cell-based systems and mouse models [4-6, 8-13]. Additionally, a limited number of less common mutant alleles of EGFR have been tested using transfection-based approaches [14-21]. Nevertheless, the biological effect of most uncommon EGFR alterations has never been evaluated. The phenotypical effect of the particular alteration detected in tumor cells may largely account for the response of the patient to treatment. In this regard, certain mutations, such as the T790M amino acid change, have been shown to confer resistance to gefitinib and erlotinib.
FUNCTIONAL ANALYSIS OF EGFR MUTANTS

FIGURE 1. Common NSCLC-associated EGFR mutations lead to increased autophosphorylation of an YFP-tagged EGFR intracellular domain. (A) Schematic representation of full-length EGFR and YFP-EGFR-ICD. The amino acid numbering includes the 24 residues of the signal peptide. YFP-EGFR-ICD contains the tyrosine kinase (TK) domain and part of the regulatory region (Reg), but lacks the extracellular and juxtamembrane (JM) domains. (B) Expression of YFP-EGFR-ICD Del746 induces morphological changes in MCF-7 cells. Unlike cells transfected with wild type (WT) YFP-EGFR-ICD, MCF-7 cells expressing YFP-EGFR-ICD Del746 frequently show long lamellipodial protrusions (arrowheads). (C) Using immunofluorescence, increased autophosphorylation of YFP-EGFR-ICD Del746 at tyrosine residues Y869 (left set of panels) and Y1092 (right set of panels) can be detected. Phosphorylation is virtually undetectable in cells expressing YFP-EGFR-ICD WT. Images were taken using 160X magnification and the exposure time indicated inside the panels. The fluorescent signal was consistently brighter using the anti-pY1092 antibody (note the shorter exposure time used). (D) Semi-quantitative comparison of YFP-EGFR-ICD autophosphorylation level using computer-assisted image analysis. Images of several transfected cells (400X magnification) were taken using 40 ms (YFP) or 160 ms (AF-594) exposure times. The fluorescence intensity in the green and the red channels was measured within a cytoplasmic area (YFP signal and AF-594 signal), and within an area outside the cells (background). In the graph, the intensity of the YFP and AF-594 fluorophores for each cell was plotted against each other using Excel, and the best-fitting trend lines (highest R²) were added. At similar expression levels (YFP intensity), the level of pY1092 is higher for YFP-EGFR-ICD bearing the Del746 mutation (white squares) than for the wild type protein (circles). The V948R mutation (open triangles) virtually abrogated autophosphorylation. The experiment was repeated twice with similar results. Graph shows the data from one experiment. au: arbitrary units.
(reviewed in [7]). Second-generation TKIs, which bind covalently to EGFR and may be active against these resistant mutants are currently being developed. To allow for a more rapid characterization of untested EGFR mutants, and to facilitate the testing of novel potential anti-EGFR agents, we aimed here to establish a simple cellular assay to evaluate the effect of EGFR mutations and the response of different EGFR variants to erlotinib. To this end, we used site-directed mutagenesis to introduce cancer-associated mutations into a YFP-tagged fragment of EGFR intracellular domain (YFP-EGFR-ICD). These chimerical proteins were transiently expressed in human cells, and the effect of their expression was assessed on a single-cell basis using immunofluorescence with phosphorylation-specific antibodies.

We demonstrate here that the YFP-EGFR-ICD-based assay can be used to evaluate the relative kinase activity and erlotinib sensitivity of EGFR mutants, and we use this approach to test several uncommon EGFR mutations.

**FIGURE 2. Activation of AKT and ERK pathways in cells expressing mutant YFP-EGFR-ICD Del746.** (A) Panels show representative images (400X) of MCF-7 cells expressing YFP-EGFR-ICD wild type, Del746 or V948R, analyzed by immunofluorescence to detect phosphorylated AKT (pAkt). Phosphorylation of endogenous AKT was only detected in cells expressing the Del746-bearing protein. (B) A similar analysis was carried out to detect phosphorylated ERK (pErk). Only cells expressing YFP-EGFR-ICD Del746 contained detectable levels of endogenous pErk. Exposure time is indicated inside the panels. DNA was counterstained with Hoechst. (C) Images (400X) illustrate two morphological characteristics of AKT phosphorylation in cells expressing YFP-EGFR-ICD Del746. On one hand, pAkt showed a preferential localization to membrane ruffles, and often accumulated at the tip of lamellipodial protrusions (arrowhead). On the other hand, cells expressing high (cell#1) or low (cell#2) levels of YFP-EGFR-ICD Del746, often contained similar levels of pAkt. A colored version of this figure can be found in appendix C – page 201.
MATERIALS AND METHODS

**Plasmid construction and site-directed mutagenesis**

In order to generate the YFP-EGFR-ICD construct, a DNA fragment encoding EGFR residues 688-1116 was amplified by PCR using primers TDG1 and TDG4, and full-length human EGFR cDNA (kindly provided by Dr. H. Nakagawa, University of Pennsylvania, Philadelphia) as template. The amplified product was digested with HindIII and KpnI and cloned into the pEYFP-C1 mammalian expression vector (Clontech, Palo Alto, CA). NSCLC-associated mutations were subsequently introduced into YFP-EGFR-ICD using the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) following manufacturer’s protocol. In all cases, the sequence of the inserts was verified by DNA sequence. The sequence of all primers used is available upon request.

**Cell culture, transfection and drug treatment**

Human breast cancer cells MCF-7 were grown in Dulbecco’s modified Eagle’s medium (BioWhittaker, Walkersville, MD), supplemented with 10% fetal calf serum (FCS), 100 units/ml penicillin, and 100 μg/ml streptomycin (Gibco-Invitrogen, Breda, The Netherlands). Cells were seeded onto sterile glass coverslips in twelve-well trays, and transfected with 0.5–1 μg of plasmid DNA using the FuGene6 transfection reagent (Roche Molecular Biochemicals, Almere, The Netherlands), following the manufacturer’s protocol. Erlotinib (Roche Pharmaceuticals, Mannheim, Germany) was added at the indicated concentration 4 hours after transfection, and the cells were incubated for 20 hours before being processed for immunofluorescence analysis. Erlotinib treatment was always performed in standard culture medium containing 10% FCS.

**Immunofluorescence and microscopy analysis**

To evaluate EGFR autophosphorylation, rabbit anti-pEGFR-Y845 (#2231, diluted 1:180) and mouse anti-pEGFR-Y1068 (#2236, diluted 1:180) antibodies were used. Note that the EGFR numbering system used by the manufacturer (Cell Signaling Technology, Danvers, MA) does not include the 24-residue signal peptide. According to the numbering system used in this report, these antibodies recognize residues pY869 and pY1092, respectively. On the other hand, rabbit anti-pAKT-S473 (#9271, diluted 1:100) and rabbit anti-pERK1/ERK2-T202/Y204 (#9101, diluted 1:100) antibodies, both from Cell Signaling Technology, were used to evaluate activation status of EGFR downstream pathways.

The immunostaining procedure was as previously described [39] with minor modifications. Briefly, cells were fixed using 3.7% formaldehyde in PBS for 30 minutes and permeabilized with 0.2% Triton X-100 in PBS for 10 minutes. Following a blocking step with 3% bovine serum...
albumin in PBS for 1 hour, the primary antibody diluted in blocking solution was applied for 1 hour. After washing with PBS, samples were incubated with Alexa Fluor 594 (AF-594)-conjugated anti-mouse or anti-rabbit secondary antibodies (Molecular Probes-Invitrogen, Breda, The Netherlands) for 45 min. Finally, the coverslips were mounted onto microscope slides with Vectashield (Vector, Burlingame, CA). The chromosome stain Hoechst 33285 (Sigma, St Louis, MO) was used to counterstain the cell nuclei. F-actin staining was carried out by incubating fixed and permeabilized cells with Rhodamine-conjugated phalloidin diluted in blocking solution for 30 minutes (Molecular Probes-Invitrogen, Breda, The Netherlands). Slides were examined using an inverted Leica DMIRB/E fluorescence microscope (Leica Heidelberg, Heidelberg, Germany). The LeicaQ500MC Quantimet software V01.01 (Leica Cambridge Ltd., Cambridge, UK) was used to collect images, keeping exposure time constant to allow for comparison of signal intensity between different samples. The same software was used to carry out semi-quantitative image analysis of YFP-EGFR-ICD expression level and pY1092 phosphorylation level. To this end, images were acquired using 400X magnification. A cytoplasmic area was selected (Figure 1D), and the intensity of the fluorescent signal within this region was measured in the green (YFP fluorophore) and the red (AF-594 fluorophore) channels. The intensity of the signal was also measured in a region outside the cell to determine background fluorescence. After subtracting the background, the fluorescence intensity of both fluorophores was plotted against each other using Excel, and the best-fitting trend line (highest R² value) was added using the “Add trend line” feature.

FIGURE 3. Evaluating erlotinib sensitivity of EGFR mutants in the context of YFP-EGFR-ICD. (A) Representative examples of MCF-7 cells expressing YFP-EGFR-ICD wild type, Del746 or Del746/T790M (green), stained for endogenous phosphorylated AKT (red). Cells were treated for 20 hours with the indicated concentration of erlotinib. The different response of each EGFR variant to erlotinib treatment is readily visualized by immunofluorescence: YFP-EGFR-ICD wild type does not induce AKT phosphorylation, and relocates into thick cytoplasmic fibrils at 10 μM erlotinib. One thousand-fold lower concentration of the drug (10 nM) inhibited Del746-induced AKT phosphorylation, and caused fibrilar relocation of the ectopic protein. The double mutant Del746/T790M did not form fibrils and induced AKT phosphorylation even in the presence of 10 μM erlotinib. (B) Images show that YFP-EGFR-ICD Del746 (green) remains phosphorylated at Y1092 (red) after relocating into fibrils in the presence of erlotinib. A colored version of this figure can be found in appendix C – page 201.
FUNCTIONAL ANALYSIS OF EGFR MUTANTS

RESULTS

Increased autophosphorylation of YFP-tagged EGFR intracellular domain the common EGFR Del746 mutation

We generated an YFP-tagged fragment of EGFR (Figure 1A) encompassing residues 688-1116 (the numbering system includes the 24 amino acid signal peptide of EGFR). This fragment, termed YFP-EGFR-ICD, contains the TK domain, as well as sequences from the adjacent regulatory motif, but lacks the extracellular and the juxtamembrane domains of the receptor. The Del746-750 mutation (hereafter called Del746) was next introduced into YFP-EGFR-ICD using site-directed mutagenesis. This mutant, representative of the common and well-characterized exon 19 deletions, was used as a positive control in the initial experiments to test the suitability of our system.

MCF-7 breast cancer cells were transfected with plasmids encoding YFP-EGFR-ICD wt or YFP-EGFR-ICD Del746, and examined using fluorescence microscopy. Approximately 30% of the cells expressing the mutant EGFR ICD showed long lamellipodial protrusions, which were not observed in cells expressing the wt fragment (Figure 1B). Cells were fixed 24 hours after transfection and immunostained using specific primary antibodies to detect phosphorylation of EGFR tyrosine residues Y869 and Y1092. Using secondary antibodies conjugated to the red fluorophore Alexa Fluor-594 (AF-594) and YFP positivity as a marker of transfection, we were able to examine YFP-EGFR-ICD phosphorylation in a single-cell basis (Figure 1C). Non-transfected MCF-7 cells did not contain detectable levels of phosphorylated EGFR. YFP-EGFR-ICD wt-transfected cells showed no or barely detectable pY869 or pY1092. In contrast, cells expressing comparable levels of the mutant protein (as indicated by the intensity of the YFP signal) showed a robust immunostaining signal for both residues. The anti-pY1092 antibody provided the clearest result, and was therefore used in subsequent analyses. Similar experiments were carried out with a shorter EGFR fragment containing only the TK domain (amino acids 688-982), but no autophosphorylation (pY869) was detected (data not shown).

We used computer-assisted image analysis to measure the intensity of the YFP and AF-594 signals in the cytoplasm of several individual cells (Figure 1D, left images). In addition to wt and Del746, an YFP-EGFR-ICD protein bearing the experimental V948R mutation was tested. This amino acid change has been shown to maintain EGFR kinase domain in an inactive conformation [22]. In line with previous data, autophosphorylation was dramatically increased by the Del746 mutation (Figure 1D, graph). It should be pointed out that cells expressing very high levels of the wt ICD showed weak but detectable Y1092 phosphorylation. The V948R change, as expected, virtually abrogated ICD autophosphorylation even at the highest levels of expression.
Altogether, these results demonstrate that the effect of NSCLC-related EGFR mutations on basal autophosphorylation can be rapidly evaluated on a single cell basis using transient transfection of YFP-EGFR-ICD and immunofluorescence.

**Activation of EGFR downstream signaling pathways in cells expressing mutant YFP-EGFR-ICD Del746**

Signal transduction downstream of EGFR and, ultimately, the cellular response to EGFR activation, relies on the integrated activity of several intracellular signaling pathways, such as those mediated by AKT or ERK. Phosphorylation of ERK and AKT at specific residues, which constitutes a key activating event in these pathways, is widely used as a marker of active EGFR downstream signaling.

MCF-7 cells transfected with YFP-EGFR-ICD wt, Del746 or V948R, were immunostained using specific antibodies to detect endogenous phosphorylated AKT (pAKT-S473) and phosphorylated ERK (pERK-T202/Y204). The levels of pAKT (Figure 2A) and pERK (Figure 2B) were undetectable in non transfected cells as well as in cells expressing YFP-EGFR-ICD wt or YFP-EGFR-ICD V948R. In contrast, phosphorylation of both AKT and ERK was clearly detected by immunofluorescence in YFP-EGFR-ICD Del746-transfected cells. A more intense signal was consistently obtained with the anti-pAKT antibody, which was therefore used in subsequent experiments. As illustrated in Figure 2C, phosphorylated AKT in YFP-EGFR-ICD Del746-transfected MCF-7 cells localized preferentially to membrane ruffles and the tip of the lamellipodial protrusions mentioned above. Remarkably, we did not observe a correlation between the expression levels of ectopic EGFR ICD and the intensity of the endogenous pAKT signal.

The sensitivity of EGFR mutants to erlotinib can be evaluated in the context of YFP-EGFR-ICD

We next evaluated the response of the Del746 mutant in the context of YFP-EGFR-ICD to the TKI inhibitor erlotinib. The TKI-resistant double mutant Del746/T790M [23, 24] was also tested. Four hours post-transfection, erlotinib at a final concentration ranging from 1 nM to 10 \( \mu \)M was added to the culture medium, and cells were incubated for 20 hours. Samples were then fixed and immunostained using the anti-pAKT antibody. As illustrated in Figure 3A, no effect of erlotinib on cells expressing YFP-EGFR-ICD wt was noted up to 1\( \mu \)M. At 10 \( \mu \)M, erlotinib induced the relocation of the chimeric protein to thick cytoplasmic filaments, reminiscent of actin cables [25]. EGFR can interact with actin [26], and the ICD fragment used in our assay includes the actin binding domain of EGFR. However, rhodamine-conjugated phalloidin failed to show co-localization with YFP-EGFR-ICD filaments (data not shown), suggesting that they do not contain actin. In cells expressing YFP-EGFR-ICD Del746, AKT phosphorylation was detected in both untreated samples and samples treated with 1 nM erlotinib. However 10 nM or higher
**FIGURE 4.** Testing the kinase activity and erlotinib sensitivity of uncommon EGFR mutants using the YFP-EGFR-ICD assay. (A) Partial amino acid sequence of EGFR exon 20 and exon 21 illustrating the location of the mutations examined (red letters). (B) Kinase activity and erlotinib sensitivity of different exon 20 mutations. Graph shows that autophosphorylation levels are lower for T790M (white diamonds) than for S768I (white squares) or Ins770SVD (black triangles). Low expression levels hampered the accurate evaluation of Ins774HV (black circles). Images show that YFP-EGFR T790M did not effectively induce phosphorylation of endogenous AKT in MCF-7 cells, and did not relocate into fibrils upon erlotinib treatment. S768I-induced pAKT was inhibited by 100 nM erlotinib and the ectopic protein relocated into fibrils at 1 μM. The phosphorylation of AKT induced by exon 20 insertions was only inhibited at 10 μM erlotinib. This drug concentration also induced relocation of YFP-EGFR-ICD Ins770SVD into fibrils. (C) Kinase activity and erlotinib sensitivity of different exon 21 mutations. Graph shows that the common L858R mutation confers higher autophosphorylation levels to YFP-EGFR-ICD than P846L and A859T. Images show that, unlike L858R, these uncommon exon 21 mutants did not induce phosphorylation of endogenous AKT. Erlotinib blocked L858R-induced pAKT at 10 nM, and caused relocation of the ectopic protein into fibrils at 100 nM. Both effects were readily abrogated by the TKI-resistant mutation T790M. In all cases, data corresponding to one experiment are shown. Each EGFR mutant was tested at least twice with similar results. A colored version of this figure can be found in appendix C – page 204.
concentration of the drug abrogated AKT phosphorylation and induced fibril formation. Surprisingly, Y1092 phosphorylation of YFP-EGFR-ICD Del746 was still detected in these fibrils (Figure 3C). In line with previous observations [24] the T790M mutation abrogated erlotinib sensitivity. Thus, AKT phosphorylation was readily detected in cells expressing YFP-EGFR-ICD Del746/T790M, even after treatment with 10 μM erlotinib.

Testing the kinase activity and erlotinib sensitivity of uncommon EGFR mutants using the YFP-EGFR-ICD assay

We next applied the YFP-EGFR-ICD-based assay to test several uncommon EGFR mutations (Figure 4A) on which limited or no biochemical information is available. These included exon 20 insertions Ins770SVD and Ins774HV, and the exon 21 point mutation P848L, which were detected during our analysis of NSCLC samples (unpublished data). We also tested the exon 21 mutation A859T identified by other groups [27-29].

These mutations were introduced into YFP-EGFR-ICD and transient transfection experiments were carried out in MCF-7 cells. The kinase activity of each mutant (autophosphorylation at Y1092 and phosphorylation of endogenous AKT), and its response to different concentrations of erlotinib were evaluated as described above. Other exon 20 and 21 mutations (T790M, S768I and L858R) that have been previously tested using transfection-based assays were also included in the assay for comparison.

Among the exon 20 mutations tested, Ins770SVD showed an intermediate level of autophosphorylation, lower than S768I, but higher than T790M (Figure 4B). It was not possible to accurately evaluate autophosphorylation of Ins774HV, since the expression level of this mutant was consistently low in all attempted experiments. AKT phosphorylation was readily detected in cells transfected with S768I, Ins770SVD or Ins774HV, but not in cells expressing T790M (Figure 4B). The phosphorylation of AKT induced by S768I was abrogated by 100nM erlotinib, and fibril formation was noted upon treatment with 1μM or higher concentration of the drug. In contrast, the phosphorylation of AKT induced by Ins770SVD or Ins774HV was only abrogated at the highest concentration of erlotinib tested (10 μM). In the case of YFP-EGFR-ICD Ins770SVD, 10 μM erlotinib also induced relocation of the chimeric protein to fibrils. No fibrils were observed in the case of YFP-EGFR-ICD T790M or Ins774HV at any of the erlotinib concentrations tested.

On the other hand, YFP-EGFR-ICD proteins bearing the uncommon exon 21 mutation P848L and A859T showed markedly lower autophosphorylation levels than YFP-EGFR-ICD L858R (Figure 4C). Neither P848L nor A859T were able to induce phosphorylation of endogenous AKT. Erlotinib induced fibrilar relocation of YFP-EGFR-ICD A859T when applied at 10 μM, but did not have any apparent effect on P848L at any of the concentrations tested. In contrast, YFP-EGFR-ICD L858R induced phosphorylation of endogenous AKT, which was inhibited by 10 nM of erlotinib, and 100
nM or higher concentrations of the drug induced fibril formation. The T790M mutation abrogated the effect of erlotinib on L858R, and the L858R/T790M double mutant readily induced AKT phosphorylation even in the presence of 10μM erlotinib.

DISCUSSION
In a subset of NSCLC patients, the presence of somatic mutations in the kinase domain of EGFR may predict the outcome of treatment with the EGFR TKIs erlotinib and gefitinib. The favorable clinical response of tumors harboring the common exon 19 deletions or the L858R mutation correlates with the high TKI sensitivity of these EGFR mutant proteins at the molecular level. However, a large variety of different EGFR mutant alleles have been identified in NSCLC patients, and it is becoming increasingly clear that different EGFR mutants may vary dramatically in their sensitivity or resistance to TKIs [14; 30-32]. This issue is of particular importance in the context of on-going prospective clinical studies in which patient selection is based on the presence of EGFR mutations. In addition, novel treatment options are being explored for those NSCLC patients bearing TKI-resistant EGFR mutations. These options include the use of second-generation irreversible EGFR TKIs currently on development, but might be extended in the future to targeting other components of the pathway.

Evaluating the biochemical characteristics of mutant EGFR proteins using in vitro or cell-based assays provides clues to the phenotypical consequences of each alteration. Several relatively uncommon EGFR mutants have been tested in transfection-based assays, using immunoblot with phosphorylation-specific antibodies to assess EGFR activity [14; 16-20]. In these studies, a homogeneous population of transfected cells was usually generated, by using viral transduction or by selecting stable transfectants, which are both labor-intensive and time-consuming procedures. We describe here a rapid cellular assay system to evaluate the kinase activity and erlotinib sensitivity of EGFR mutants, using an YFP-tagged fragment of EGFR intracellular domain (YFP-EGFR-ICD) and immunofluorescence. Our assay presents several advantageous characteristics with respect to previously used methods. First, by using a fragment of the receptor lacking the extracellular domain one would expect to reduce interference from the experimental context, which may have been partially responsible for some controversial findings [4, 6, 9, 18, 19]. Furthermore, the use of a shorter EGFR fragment instead of the full-length receptor renders the mutagenesis procedure more efficient. Finally, by evaluating EGFR activity in a single-cell basis, our assay circumvents the need of a homogeneous population of transfected cells allowing the use of transient transfection. The whole procedure of testing a new EGFR mutant, including site-directed mutagenesis (2 days), verification of the construct by sequencing (1 day), transfection, immunostaining and scoring (3 days) can be completed in approximately
one week. In comparison, the generation of a population of stably transfected cells would typically require several weeks of selection in antibiotic-containing medium.

By using the well-characterized EGFR mutant Del746 as control, we demonstrate that the YFP-EGFR-ICD-based assay readily identifies differences between this mutant and the wild type protein in terms of autophosphorylation, activation of downstream signaling pathways and sensitivity to erlotinib. We also show that computer-assisted measurement of fluorescence intensity can be used to obtain a semi-quantitative comparison of autophosphorylation levels between different mutants. It must be acknowledged that immunofluorescence staining is a less quantitative approach than immunoblot to evaluate protein phosphorylation levels, a disadvantage that is, in our view, counterbalanced by the preservation of cellular morphology. Morphological examination allowed us to observe that phosphorylated endogenous AKT preferentially localizes to membrane ruffles and the tip of lamellipodial protrusions in cells expressing mutant EGFR-ICDs. This observation is consistent with the localization of activated endogenous AKT in growth factor-stimulated cells and its role in cell motility ([33], and references therein).

Microscopy analysis led to the unexpected observation that erlotinib treatment induces the relocation of the YFP-EGFR-ICD chimeric protein to thick cytoplasmic filaments. Importantly, the wild type protein formed fibrils only at 10 μM erlotinib, whereas the TKI-sensitive mutants relocated to fibrils in the presence of 10-100 nM erlotinib, and this effect was fully abrogated by the erlotinib-resistant T790M mutation. These observations suggest that the relocation of the chimeric YFP-EGFR-ICD protein into thick fibrils at lower drug concentrations is a marker of erlotinib sensitivity in our assay. The molecular basis for this effect of erlotinib is presently unclear. Additional experiments are required, for example, to clarify why YFP-EGFR-ICD Del746 remains phosphorylated (pY1092) in these fibrils in the presence of 500 nM erlotinib, even if downstream signaling (pAKT) is inhibited. We speculate that erlotinib binding to the ATP-binding site in the context of YFP-EGFR-ICD molecule may introduce a conformational alteration sufficient to lead to the aggregation of the chimeric protein. Importantly, we have noted that such effect is not erlotinib-specific, since a similar relocation can be induced by gefitinib treatment (data not shown). It remains to be further examined if fibril formation is a general effect of TKI-mediated EGFR inhibition. In this case, the YFP-EGFR-ICD assay system could be adapted for high content screening of potential anti-EGFR agents, since the shift from a diffuse YFP signal to a fluorescent signal concentrated in thick fibrils would be readily detected using automated image analysis. This possibility is particularly appealing in the context of the currently on-going effort to develop second-generation irreversible TKIs and other agents that may circumvent resistance to TKI associated with the presence of certain types of EGFR mutations.
Over the last three years, we have carried out mutational analysis of EGFR in tumor samples from nearly 300 NSCLC patients [34, 35] (and unpublished data). Several uncommon EGFR mutants identified in the course of this analysis were tested using the YFP-EGFR-ICD assay. Mutations in exon 20, most notably T790M, are usually associated with resistance to gefitinib and erlotinib (reviewed in [36]). We found that Ins770SVD and Ins774HV are more resistant to erlotinib than S768I in our assay. These results are in line with previous data showing increased erlotinib resistance for a similar (Ins770NPG) mutant [14]. Furthermore, our results show that both insertions confer higher kinase activity than T790M, thus underscoring an important difference between these two types of exon 20 alterations. In this regard, there has been some controversy regarding the enhanced kinase activity of T790M-mutant EGFR [21]. We noted that the autophosphorylation level (pY1092) of YFP-EGFR-ICD T790M was indeed higher than that of the wild type protein at lower expression levels (compare graphs in Figure 1D and Figure 4B), as reported by Vikis et al [21]. However, the ability of this mutant to activate downstream signaling was clearly reduced in comparison to the ICD constructs bearing the Del746, L858R, S768I, Ins770SVD or Ins774HV mutations. Recent structural analyses indicate that enhanced activity of EGFR mutants may derive form the disruption of autoinhibitory interactions that suppress EGFR basal activity [22, 37]. Our data suggest that exon 20 insertions may disrupt these interactions to a greater extent than the T790M point mutation.

On the other hand, our functional analysis indicates that the uncommon exon 21 mutation P848L is not a kinase-activating mutation and does not confer increased sensitivity to erlotinib. This change has been detected in both tumor and normal tissues from NSCLC patients [38; Gallegos-Ruiz et al., unpublished data]. Furthermore, although EGFR and K-ras mutations are in general mutually exclusive in NSCLC patients [36], a K-Ras mutation (G12V) was detected in tumor cells bearing the P848L allele (Gallegos-Ruiz et al., unpublished). Like P848L, the A859T variant, which has been detected in two non-responding NSCLC patients [27, 28], did not confer increased kinase activity or erlotinib sensitivity in our test. Together, the results from the YFP-EGFR-ICD assay and the clinical behavior of tumors bearing these alterations suggest that P848L and A859T are likely to be uncommon, functionally silent EGFR polymorphisms.

CONCLUSION

In conclusion, we describe here the use of a simple cellular assay that can be easily implemented to functionally evaluate EGFR variants. The ability to rapidly obtain functional information on EGFR variants of unknown relevance might prove important in the future for the management of NSCLC patients bearing uncommon EGFR mutations. In addition, our assay may be used to determine the response of resistant EGFR mutants to novel second-generation TKIs or to other therapeutic agents targeting the EGFR signaling pathway.
REFERENCES


ACKNOWLEDGEMENTS

We are very grateful to Dr. H. Nakagawa, University of Pennsylvania, Philadelphia, for providing full-length EGFR cDNA and Cora Beckers, VUMC, for assistance with the phallolidin staining. This work was partially supported by Roche.
COMBINED ASSESSMENT OF EGFR PATHWAY-RELATED MOLECULAR MARKERS AND PROGNOSIS OF NSCLC PATIENTS

Mariëlle I Gallegos Ruiz, Karijn Floor, José Ferreira, Katrien Grünberg, Frederik BJM Thunnissen, Jeroen AM Belien, Gerrit A Meijer, Godefridus J Peters, Egbert F Smit, José A Rodriguez and Giuseppe Giaccone

IN PREPARATION
ABSTRACT

Purpose: To evaluate the prognostic value of the combined assessment of multiple molecular markers related to the EGFR pathway in resected NSCLC patients.

Experimental Design: Tumor specimens of 178 NSCLC patients were collected and analyzed for EGFR and KRAS mutation status by DNA sequencing, and for EGFR copy number by FISH. Tissue microarrays were generated and used to determine the expression of various downstream proteins by immunohistochemistry. We analyzed the association between the different markers and the correlation with patient prognosis.

Results: We used a random survival forest predictor to assess the prognostic value of the EGFR related markers when analyzed in combination. Analyses were performed on both the total patient group (n=178) and a subgroup of selected patients with pathological stage I and II that underwent an R0 resection and had no significant co-morbidity (n=65). The use of EGFR pathway-related markers provides improved prognostic information over disease stage for both the total patient group (1-Harrell’s index error rate 0.446 vs 0.483) and for the selected patient group (1-Harrell’s index error rate 0.449 vs 0.499). Cytoplasmic PTEN, cytoplasmic pSTAT5, cytoplasmic HIF1α, and nuclear pCMET (residue 1349) are the most important markers in this prognostic model common to both groups.

Conclusions: The EGFR-related marker set evaluated here has a 7-10% greater prognostic value than the commonly used pTNM staging. A combined immunohistochemical analysis of several of these markers, namely PTEN, pSTAT5, HIF1α and pCMET, may help refine prognosis in NSCLC patients.
INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide [2], and non-small cell lung cancer (NSCLC) represents 85% of lung tumors. The Epidermal Growth Factor Receptor (EGFR) pathway plays a fundamental role in the carcinogenesis and progression of various tumor types, including NSCLC [3]. EGFR (ErbB-1) is a member of the ErbB family of receptors, which also includes HER2/neu (ErbB-2), Her 3 (ErbB-3) and Her 4 (ErbB-4). Autophosphorylation of EGFR intracellular tyrosine kinase domain results in activation of several downstream signaling pathways, including the PI3K, STAT and the MAPK pathways, which regulate biological responses such as proliferation, cell motility, angiogenesis, cell survival and differentiation [4,5].

An improved understanding of EGFR signaling has led to the development of anticancer therapeutics directed against EGFR, including the tyrosine kinase inhibitors (TKIs) gefitinib and erlotinib [6]. Objective responses to these agents occur only in a small subset of unselected NSCLC patients, and several molecules involved in EGFR signaling have been evaluated in an effort to identify markers of TKI sensitivity. Such molecular markers include specific mutations in EGFR or K-ras, EGFR gene copy number [6-10], the activation status of AKT and STAT signaling pathways [11] and the expression level of HER2 [12,13]. More recently, amplification of the CMET receptor [14], the expression of epithelial to mesenchymal transition (EMT) markers, such as E-cadherin and vimentin [15,16], and the downregulation of HIF1 have also been linked to responsiveness to EGFR targeted agents [17].

Although several of these markers have been identified as potential predictors for response to EGFR tyrosine kinase inhibitors in patients with advanced NSCLC, some of them have also been shown to be prognostic for survival, irrespective of treatment. To be able to distinguish between these two effects is important.

The presence EGFR mutations has been proposed to be a positive prognostic factor [18], whereas high EGFR copy number and the presence of KRAS mutations have both been linked to poor prognosis in resected NSCLC patients [19-21]. Several other markers have also been associated to poor prognosis in NSCLC (EGFR, CMET, E-CADHERIN, pAKT [22-26], but at present there is no single marker that can be used to guide therapy or predict prognosis of NSCLC patients.

Given the importance of the EGFR pathway in NSCLC, we hypothesized that the combined analysis of several of these molecular markers (Figure 1), which provide information on the activity/sensitivity of the EGFR signaling pathway at different points, might be related to the prognosis of NSCLC patients when analyzed in combination. This analysis may help to distinguish the prognostic implication of the EGFR pathway and from its predictive value in patients treated with agents targeted to this pathway. Thus, we carried out an analysis of EGFR and KRAS.
mutational status, EGFR copy number, and the expression of EGFR, HER2, pCMET, pAKT, PTEN, pSTAT3, pSTAT5, pERK, HIF1α, E-CADHERIN and VIMENTIN, in resected NSCLC patients.

**FIGURE 1. Overview of markers analyzed in this study.** A graphic display of a selection of EGFR pathway related markers. The marker analyzed in this study are indicated with an asterisk (*). Dashed lines indicate hypothesized interactions.

**MATERIALS AND METHODS**

**Patients and samples**

Radically resected tumor specimens of 178 NSCLC patients were collected. For 148 patients, both frozen and paraffin-embedded tissue was available; for 30 patients only paraffin-embedded material was available. Samples were obtained from patients with pathological stage I, II or III and 24% of patients received (neo-)adjuvant chemo/radiotherapy. A full description of patient characteristics is provided in Table 1. The study was carried out in accordance with the ethical guidelines of our institution concerning informed consent about the use of patient’s material after surgical procedures.
Isolation of genomic DNA

DNA was isolated from frozen tissue (n=148). Sections of tissue samples flanking those used for DNA isolation were verified by the study pathologists (KG and ET) to contain at least 50% of tumor cells. Genomic DNA was extracted from frozen samples using Trizol, following manufacturer instructions (Life Technologies, Breda, The Netherlands).

PCR amplification and DNA sequencing

Mutation analysis was carried out on 148 patients for which frozen tissue samples were available, because paraffin-embedded samples might yield a higher proportion of false positive results [27,28], and we avoided using two different sources of samples (frozen and paraffin) for one type of analysis. We used 100 ng of genomic DNA derived from tumor cells as template in nested PCR reactions to amplify DNA fragments corresponding to exons 18-21 of EGFR, and exons 1 and

| Table 1. Clinicopathological characteristics of the patients included in this study |
|-------------------------------------------------|-----------------|-----------------|
| Characteristic                                  | Unselected group | Selected group  |
|                                                 | n=178 n (%)      | n=65 n (%)      |
| Gender                                          |                 |                 |
| Male                                            | 127 (71)        | 47 (72)         |
| Female                                          | 51 (29)         | 18 (28)         |
| Histology                                       |                 |                 |
| Adenocarcinoma                                  | 64 (36)         | 25 (39)         |
| BAC                                             | 6 (3)           | 2 (3)           |
| Squamous Cell Carcinoma                         | 77 (43)         | 27 (42)         |
| Large Cell Carcinoma                            | 24 (14)         | 9 (14)          |
| Others                                          | 7 (4)           | 2 (3)           |
| Smoking status                                  |                 |                 |
| Never                                           | 3 (2)           | 1 (2)           |
| Former                                          | 70 (39)         | 26 (40)         |
| Current                                         | 60 (34)         | 19 (29)         |
| Unknown                                         | 45 (25)         | 19 (29)         |
| Tumor stage                                     |                 |                 |
| I                                               | 82 (46)         | 47 (72)         |
| II                                              | 47 (26)         | 18 (28)         |
| III                                             | 33 (19)         | 0 (0)           |
| Unknown                                         | 16 (9)          | 0 (0)           |
| Treatment                                       |                 |                 |
| No treatment                                    | 133 (75)        | 53 (82)         |
| pre operative chemo                             | 28 (16)         | 10 (15)         |
| post operative chemo                            | 3 (2)           | 2 (3)           |
| post operative radio                            | 11 (6)          | 0 (0)           |
| pre operative chemo-radio                       | 3 (1)           | 0 (0)           |
| Resection                                       |                 |                 |
| R0                                              | 123 (69)        | 65 (100)        |
| R1                                              | 36 (20)         | 0 (0)           |
| R2                                              | 9 (5)           | 0 (0)           |
| Unknown                                         | 10 (6)          | 0 (0)           |

| R0: Complete resection, R1: Microscopic residue, R2: Macroscopic residue: Unknown: uncertain |
2 of KRAS. The PCR protocol and the sets of primers have been described in detail previously [29]. PCR products were purified using a presequencing kit (Amersham Biosciences, Roosendaal, The Netherlands) and sequenced with both forward and reverse primers using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA), using the ABI PRISM™ 3100 Genetic analyzer (Applied Biosystems). Mutations were confirmed by sequencing independent PCR products.

**Fluorescent In Situ Hybridization (FISH)**

Fluorescent In Situ Hybridization (FISH) was performed only on frozen sections since this analysis provides poorer results on paraffin-embedded tissue [30]. 4 μm frozen sections were fixed with methanol/acetic acid (3:1) and pre-treated by digestion with 0.01% pepsin/0.2 N HCl at 37°C for 2 min, and incubated for 10 min in 50 mM MgCl2/PBS followed by 10 min in 50 mM MgCl2/3.7% formaldehyde/PBS. After 2 h incubation with 70% formamide/0.6x SSC, the sections were dehydrated with alcohol. Following pretreatment, 10-15 μl LSI EGFR Spectrum Orange/CEP7 Spectrum Green probe (Vysis, Abbot Laboratories, Downers Grove, IL, USA) was applied, the section was covered with a coverslip, and sealed with rubber cement. Following a denaturation step at 80°C for 10 min, slides were placed in a humidified chamber at 37°C for 20-24h. Then, sections were washed with 1.5 M urea/0.1x SSC at 45°C for 30 min, and with 2x SSC for 2 min. Finally, sections were counterstained with 4',6-Diamidino-2-phenylindole (DAPI, Sanbio BV, Uden, The Netherlands), dehydrated with alcohol, air-dried and mounted using Vectashield (Brunschwig Chemie, Amsterdam, The Netherlands).

**Scoring of FISH analysis results**

FISH slides were evaluated using a Leica DMRA fluorescent microscope (Leica Microsystems BV, Wetzlar, Germany) with a 60x PL Fluotar oil immersion objective (NA=1.40). Scoring was done by two independent observers (KF and MGR). For every sample, the complete section was screened for homo/heterogeneity of the FISH signals. The signals in 200 tumor cells were counted in at least three representative microscopic fields. The number of cells having 0, 1, 2, 3, 4, 5 or ≥6 red signals or clusters was noted and samples were categorized as described previously [31]. Samples were considered as having high EGFR polysomy when ≥4 dots per nucleus were present in ≥40% of tumor cells, and as having EGFR amplification when tight EGFR gene clusters were present in ≥10% of cells.

**Tissue Micro-Array (TMA) construction**

Paraffin-embedded tumor material of 178 patients was cut into 4 μm-thick sections and placed onto glass slides. Slides were stained with hematoxylin and eosin and a pathologist (KG) verified
the presence of tumor cells and marked the tumor area. 0.6 mm diameter biopsies were taken from the donor block, two from the tumor and one from the normal tissue area surrounding the tumor. Biopsies from the donor blocks were included in recipient tissue array blocks using a precision tissue array instrument (Beecher Instruments, Sun Prairie, WI, USA).

**Immunohistochemistry**

TMA sections were deparaffinized using xylene and dehydrated in alcohol. To block the endogenous peroxidase activity, tissue slides were incubated in methanol/0.3% H2O2 for 30 minutes. Antigen retrieval was carried out by heating the slides in 0.1 M Sodium Citrate or 1 mM Tris/EDTA (pH 9.0) for 30 minutes. Sections were then incubated with the primary antibody overnight at 4°C, using sections incubated with antibody diluent (Immunologic, Duiven, The Netherlands) as negative control. Sections were developed using the DAKO EnvisionTM visualization system (Dakocytomation, Heverlee, Belgium). Pre-treatment conditions and antibody dilutions are indicated in Table 2. Note that the phosphorylation status of cMET and AKT at several residues was analyzed using different phospho-specific antibodies.

**Immunohistochemistry Scoring**

Protein expression determined by immunohistochemistry was evaluated using an Olympus BX50F bright field microscope (Olympus Optical Co Ltd, Tokyo, Japan) with a 40x Plan objective (NA=0.65). Scoring was done by two observers (KF and MGR). For each protein, intensity (negative: 0, weak positive: 1, moderately positive: 2, strong positive: 3) and percentage of positively stained cells were scored. Some cases suffered tissue loss or lack of tumor cell representation to an extent that precluded the evaluation of protein expression. The different subcellular localization of the proteins was recorded as nuclear (N), membrane-associated (M) or cytoplasmic (C) [32]. For example, a marker described as pCMET.1003.N refers to CMET phosphorylated at residue 1003 localized in the nucleus. We also note that, in the particular case of pERK, some samples showed specific pERK granules in the cytoplasm and were categorized as pERK.Gr. The staining intensity value was multiplied by the percentage of positive cells [33], yielding a final expression score ranging from 0 to 300.
Statistics

The ’hist’ function of S-PLUS was used to stratify the protein expression data into a recommended number of two to four categories, the first category representing ‘no expression’ [34]. The stratification smoothes the discretely but very sparsely distributed protein scores, making it possible to use non-parametric methods to test associations between proteins and other variables. Associations between various markers were tested using Fisher’s exact test. Multiple testing was taken into account when drawing conclusions from the p-values of the tests, by means of bounds on false discovery rates [35]. Random survival forest (RSF) predictors of patient survival time were constructed with the RSF software [36]. In short, given a training data set of survival/censoring times, censoring status and other variables measured on a sample of patients, the RSF software constructs an ensemble of survival function estimates - the predictor. Given the values of the covariates of a patient not in the training data set, RSF associates a survival function of the predictor to them, and this survival function can be used to predict the survival time of a ‘new’ patient. The quality of RSF predictors is assessed in terms of an error rate which takes values between 0 (if the prediction is 100% accurate) and 1 (if the prediction is 100% inaccurate). An error rate of 0.5 indicates that the prediction is as accurate as ‘random guessing’. This error rate is defined as one minus Harrell’s concordance index [37]. A more detailed explanation of the RSF predictors can be found as supplementary information. We have used a ‘random subsampling method’ to estimate and compare the mean error rates of several predictors based on different subsets of markers and patients. The number of subsamples drawn (1000) is large enough to guarantee the statistical significance of the differences found between mean error rates. The log-rank test was used to test associations between survival and protein markers.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Antibody</th>
<th>Type</th>
<th>Source</th>
<th>Antigen retrieval Method</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td>EGFR pharm DX kit K1492</td>
<td>monoclonal</td>
<td>DakoCytomation, Heverlee, Belgium</td>
<td>Prot K proteolytic digestion</td>
<td>1:1</td>
</tr>
<tr>
<td>Her2</td>
<td>Her-2 non commercial</td>
<td>monoclonal</td>
<td>Dr M vd Vijver, Avl, Amsterdam, The Netherlands</td>
<td>10 mmol/L of citrate buffer (pH 6.0), microwave</td>
<td>1:100</td>
</tr>
<tr>
<td>Cmet</td>
<td>cMet pY1003</td>
<td>polyclonal</td>
<td>Invitrogen Biosource, Breda, The Netherlands</td>
<td>1mM Tris/EDTA (pH 8.0), microwave</td>
<td>1:1000</td>
</tr>
<tr>
<td>Cmet</td>
<td>cMet pY1349</td>
<td>polyclonal</td>
<td>Invitrogen Biosource, Breda, The Netherlands</td>
<td>1mM Tris/EDTA (pH 8.0), microwave</td>
<td>1:1000</td>
</tr>
<tr>
<td>PTEN</td>
<td>PTEN 138G6, #9559</td>
<td>monoclonal</td>
<td>Cell signaling, Boston, MA, USA</td>
<td>10 mmol/L of citrate buffer (pH 6.0), autoclave</td>
<td>1:200</td>
</tr>
<tr>
<td>pAkt</td>
<td>pAkt ser473, #9277</td>
<td>monoclonal</td>
<td>Cell signaling, Boston, MA, USA</td>
<td>11 mmol/L of citrate buffer (pH 6.0), microwave</td>
<td>1:150</td>
</tr>
<tr>
<td>pAkt</td>
<td>pAkt thr309 (244F9) #4056P</td>
<td>monoclonal</td>
<td>Cell signaling, Boston, MA, USA</td>
<td>1mM Tris/EDTA (pH 9.0), microwave</td>
<td>1:150</td>
</tr>
<tr>
<td>pStat3</td>
<td>pSTAT3 tyr705 (D3A7)</td>
<td>monoclonal</td>
<td>Cell signaling, Boston, MA, USA</td>
<td>1mM Tris/EDTA (pH 8.0), microwave</td>
<td>1:50</td>
</tr>
<tr>
<td>pStat5</td>
<td>pSTAT5a/b tyr 694/99</td>
<td>monoclonal</td>
<td>Advantex Bio Reagents, Conroe, TX, USA</td>
<td>10 mmol/L of citrate buffer (pH 6.0), autoclave</td>
<td>1:500</td>
</tr>
<tr>
<td>pErk</td>
<td>p44/42 MAPK, thr 202/tyr204 (E10) #9106</td>
<td>monoclonal</td>
<td>Cell signaling, Boston, MA, USA</td>
<td>10 mmol/L of citrate buffer (pH 6.0), autoclave</td>
<td>1:100</td>
</tr>
<tr>
<td>HIF1α</td>
<td>HIF1α, clone 54, 610959</td>
<td>monoclonal</td>
<td>BD Bioscience, San Jose, CA, USA</td>
<td>Target retrieval solution, microwave</td>
<td>1:500</td>
</tr>
<tr>
<td>E-cadherin</td>
<td>E cadherin</td>
<td>monoclonal</td>
<td>SigmaAldrich, St Louis, MO, USA</td>
<td>Target retrieval solution, microwave</td>
<td>1:100</td>
</tr>
<tr>
<td>Vimentin</td>
<td>Vimentin, Clone V9, #3400</td>
<td>monoclonal</td>
<td>SigmaAldrich, St Louis, MO, USA</td>
<td>1mM Tris/EDTA (pH 8.0), microwave</td>
<td>1:100</td>
</tr>
</tbody>
</table>
RESULTS

EGFR, K-Ras mutations and EGFR copy number

EGFR mutation analysis was successful in 136 patients and mutations were identified in 4% (n=5) of the patients. Mutations observed were P848L (n=1), L858R (n=1), DelL746-752S (n=1), the double mutation E709K+L858R (n=1) and the double mutation S768I+L861Q (n=1). EGFR mutation P848L was not considered to be a cancer-specific EGFR mutation based on previous findings by us and others [38,39]. K-ras mutations were observed in 18% (n=25) of the 139 patients whose samples could be analyzed. Mutations were predominantly observed in codon 12 (n=19) but were also detected in codon 13 (n=3) and codon 61 (n=3). Finally, FISH could be evaluated in 138 patients. Amplification (tight EGFR signal clusters) was observed in 6% of the patients, whereas high polysomy, defined as more than 4 EGFR signals per cell in more than 40% of cells, was observed in 8% of the patients.

The low prevalence of EGFR mutations in our patient cohort (4%) was similar to other European studies [40,41]. We did not observe an overlap of KRAS mutations and EGFR mutations in the same patients, in line with previous studies [18,20]. An association between EGFR mutations and high EGFR copy number has been described previously [42]. Out of 5 patients carrying an EGFR mutation, 2 patients had high EGFR copy number and 1 patient had EGFR amplification.

Protein markers

The expression, phosphorylation status, and subcellular localization of proteins evaluated in 178 patients using the TMA by IHC are referred to as protein markers. As described in the Material and Methods section, for each protein we scored the staining intensity and the percentage of positively stained cells. In Figure 2, representative images of immunohistochemical staining for EGFR (A), pCMET (B-C), E-CADHERIN (D), PTEN (E) and pSTAT3 (F) are shown.

Selection of patient population

We anticipated that patient heterogeneity might influence the outcome of survival and association analyses in the whole patient group (n=178). There were patients in this group who had either microscopic (R1, 20% of patients) or macroscopic (R2, 5% of patients) residual disease after resection, and the type of resection was uncertain in 6% of the patients. In addition, some patients had significant co-morbidity, such as other tumors in addition to their NSCLC (29% of patients). Finally, 19% of patients had pathological stage III. All the above factors are strong negative prognostic factors. In order to define a more homogeneous patient population, in which the potential role of biological markers could be more easily identified, we selected those patients who had a complete resection (R0), no co-morbidity, and pathological
stage I or II. A subgroup of 65 patients (referred to as "selected patient group") met these criteria. A full description of the characteristics of the total and selected patient groups is provided in Table 1.

**Association between markers in the selected patient group**

The association analyses described below were initially performed on both the unselected and selected patient groups. We describe here the results obtained on the selected patient group since the analysis of the unselected patient group yielded weaker results than the analysis of the selected group. As described in the Materials and Methods section, we have categorized the data in order to smooth the sparsely distributed protein scores, and to allow the use of non-parametric methods to test associations between proteins and other variables. All analyses were performed on two, three and four categories. We focus here on the results obtained with two categories, because taking three or four categories, yielded overall consistent results, but diluted the power of the tests. Fisher’s exact test was used to test the pairwise association between markers.
Regarding the association between DNA and protein markers, evidence was found of an association of EGFR FISH with cytoplasmic EGFR and with cytoplasmic pCMET.1003 (P=0.002 and P=0.004, respectively). By calling both associations significant we incur in a false discovery rate of at most 8%. Evidence was also found of an association between KRAS mutations and a granular staining pattern of pERK (P=0.009, false discovery rate of 30% or less), see supplementary Figure S1 for details on this association. Finally, more than 60 protein-protein associations were found (at a false discovery rate of 10% or less), which is consistent with the known biological and functional relationship between these molecules.

Prognostic value of EGFR related markers in NSCLC patients

In order to assess the prognostic value of the various markers analyzed we constructed random survival forest (RSF) predictors with both the unselected and selected patient groups, and quantified their prediction value. In order to construct the predictors, we stratified the protein data into an average number of four categories per protein marker, as indicated in Materials and Methods; using fewer or more categories (or using the raw data without stratification) generally yielded poorer results (not shown).

We started by constructing a ‘baseline predictor’ based on pTNM stage alone, since this is currently regarded as the most reliable prognostic factor for NSCLC [43]. Next, two other predictors were constructed, one based on the combination of clinical variables (gender, histology, smoking status, stage, treatment, type of resection) with DNA-protein markers, and the other based only on DNA-protein markers (without clinical variables). The performance of each predictor was assessed in terms of the mean error rate incurred by using a randomly drawn two-thirds of the data to predict the remaining one third. More precisely, estimates of the mean and standard deviation of the error rate (defined in Materials and Methods) were obtained by repeating 1000 times the procedure of constructing the RSF predictor with one portion of the data and using it to predict the other portion. The baseline predictor, based on stage alone, had a mean error rate of 0.483 and a standard deviation of 0.053 when used on the unselected group, and a mean error rate of 0.499 and a standard deviation of 0.094 when used on the selected group. The predictor constructed with the clinical variables plus the markers had an average error of 0.451 when used on the unselected group and of 0.465 when used on the selected group, the standard deviations being 0.063 and 0.106 respectively. The average error of the predictor that combines the protein and gene markers (leaving out the clinical information) was 0.446 when used on the unselected group and 0.449 when used on the selected group, the corresponding standard deviations being 0.062 and 0.109. The average error rates incurred using the various input variables are summarized in Table 3. The difference in predictive power between the stage-based predictor and the DNA-protein marker-based predictor is 7% for the
selected group and 10% for the selected patient group. Since these figures are based on 1000 independent repetitions of the same procedure, they are highly significant. These data, therefore, show that using DNA and protein markers improves prediction relative to using pTNM staging alone. The fact that using all variables yields somewhat worse predictions as compared to using just all DNA and protein markers might be explained by the fact that the number of missing or ‘non-informative’ data increases with the inclusion of more variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unselected group n=178</th>
<th>Selected Group n=65</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTNM</td>
<td>0.483 0.053</td>
<td>0.499 0.094</td>
</tr>
<tr>
<td>Clinical and DNA-protein markers</td>
<td>0.451 0.063</td>
<td>0.465 0.106</td>
</tr>
<tr>
<td>DNA-protein markers</td>
<td>0.446 0.062</td>
<td>0.449 0.109</td>
</tr>
</tbody>
</table>

The RSF algorithm also provides information about which variables play a more important role in the prediction by means of ‘variable importance’ analysis [36]. Figure 3 shows scatter plots of the error incurred versus the variable ranked first in terms of importance in the 1000 runs of our ‘training plus predicting’ procedure. In each of these plots, there are certain markers that are more often ranked first in importance. Thus, pSTAT5.C and HIF1α.C appear as prominent variables when using both unselected and selected patient groups. In contrast, both VIMENTIN.C and E-CADHERIN.M play an important role in the unselected patient group but they are almost absent in the selected group, whereas the converse was observed for pAKT473.C and HER2.C. Interestingly, none of the DNA markers was particularly useful for predicting survival. Among the clinical parameters, the most important variables were stage (in both groups), (neo-)adjuvant treatment (in the unselected group), and histology (in the selected group).

The markers that overlapped, ranking first in both patient groups, were pSTAT5.C, HIF1α.C, PTEN.C and pMET.1349.N. We assessed the effect that these markers have on survival, individually as well as in combination. Figure 4 shows that the expression of each of the first three proteins and the absence of pMET.1349.N are associated with increased survival. To examine the cumulative effect of the four markers we compared the survival curves of patients having one or two markers of increased survival with those of patients having three or four markers of increased survival. As shown in Figure 4, the separation of survival curves in terms of the combination of markers is significantly improved relative to that achieved with the markers individually.
DISCUSSION

At present, the pTNM stage remains the most reliable prognostic factor for NSCLC [43]. However, staging alone is unable to correctly predict survival in a significant proportion of patients who undergo radical resection. The use of biological markers has been investigated as a way to increase the ability to formulate prognosis in patients.

Much effort is currently ongoing to identify biomarkers that are prognostic for survival, and also markers that are predictive of response to systemic therapies. cDNA microarray technology has been widely used to this end, but this technique is still far from clinical implementation mainly due to the need for validation and standardization across laboratories and also the high costs. Immunohistochemistry (IHC), on the other hand, is a widely accepted technique of assessing protein expression with much lower costs than microarrays, although validation and standardization are also critical issues. Besides the semi quantitative measurement of expression levels, IHC provides information on subcellular localization of the protein under study.

Given the importance of EGFR-mediated signaling in NSCLC, various EGFR-related proteins have already been studied using immunohistochemistry in NSCLC tumor samples. Several markers...
have shown to provide some prognostic information, such as E-CADHERIN, EGFR [22,23] and pAKT [24], although none of them has proven to be sufficiently useful in clinical diagnostics in terms of prediction of response to treatment or prognosis. Since simultaneous analysis of various markers could potentially increase prognostic significance over individual markers, we have used a multi-marker approach, using Tissue Micro Arrays (TMAs) and random forest survival predictors, to investigate the relevance of various EGFR pathway-related markers, and their association with NSCLC patient prognosis.

It is known that prognostic markers based on IHC can provide inconsistent or contradictory results, due to the use of different antibodies and processing methods [44,45] as well as different scoring and categorization systems. All these issues emphasize the need for standardized processing and scoring procedures as means of guaranteeing successful results. In addition, since quantitation is subjective and subcellular localization may matter [32], it would be desirable that IHC findings be reported carefully and in detail. In this study, we have used resection specimens to perform IHC on TMAs to identify protein expression, and we have scored the staining intensity, the percentage of positive cells, and the subcellular localization of proteins.

In order to study the association between the different markers analyzed, we selected a more homogeneous population within the total patient group. The selection was based strictly on clinical criteria (stage, type of resection and co-morbidity), and was motivated by our preliminary observation that considering the more heterogeneous total patient group comprising patients with other types of cancer, all tumor stages, and no complete resections, would weaken the power of detecting associations.

Mutation analysis of EGFR and KRAS is widely being evaluated for its ability to guide therapy of NSCLC patients [10,20,41,46]. We investigated whether mutation status was correlated with particular 'protein-profiles' in order to get insight into the biological interactions between various molecules and different genotypes of patients. We did not observe a statistically significant association of EGFR mutations with EGFR copy number or any of the other markers analyzed, perhaps due to the low number of patients carrying an EGFR mutation (4%) in our patient cohort. We found, however, a significant correlation between the presence of KRAS mutations and a granular staining pattern of pERK. The molecular basis for such correlation remains to be elucidated. On the other hand, we found that high EGFR copy number is significantly related with higher cytoplasmic expression of EGFR protein. This suggests that increased EGFR gene copy number might result in overexpression of the receptor in the cytoplasm, rather than on the membrane, perhaps related to abnormal post translational modifications. In addition, high EGFR copy number was related with higher pCMET.1003 in the cytoplasm, providing evidence for a relation between EGFR copy number and CMET activation. Interestingly, it has been recently reported that EGFR is necessary for CMET activation in mammary carcinoma cells [47].
The prognostic value of the markers analyzed was investigated by using RFS. The random forests predictors used here have good predictive performance relative to other predictors, account for linear or non-linear interactions between covariates, are invariant to monotonic transformations of the covariates and robust against overfitting [36,48]. Thus, they are well-suited to deal with prediction problems based on large numbers of covariates and features of random forests to predict survival have been used in previous studies [49,50]. The RSF analysis also provides a measure of 'variable importance', ranking the variables in terms of their usefulness in prediction. Finally, although random forests are able to impute missing data, our results have shown that discarding the missing data gives the best results in our case.

We first computed a baseline prediction error rate based on stage alone and compared this to error rates obtained by adding clinical and DNA-protein variables. When the predictor was constructed with all the protein and DNA markers, the average error was decreased by 7-10%, indicating that taking multiple markers into account does improve prediction relative to using
pTNM staging alone. The improvement may, in fact, be greater than it appears from our calculations, because there are a number of patients with missing data on the protein level, due to tissue loss or lack of tumor cell representation.

The prognostic value of the markers was evaluated on both the unselected and selected group, in order to compare their predictive power in the different groups, and to get an indication of which markers are important in both groups. The fact that, the predictor based on the selected group perform as well as those based on the unselected group (despite the difference in sample sizes), seems to support the idea that focusing on subgroups of patients selected on the basis of clinical information could improve prognostic power. Thus, focusing on specific patient subsets might be worthwhile in the search for novel DNA and protein markers.

Those markers that overlapped in importance in both groups when using our prognostic model (PTEN.C, pSTAT5.C, HIF1α.C and pCMET.1349.N) did provide prognostic information when analyzed in combination, although each of these markers did not have a great prognostic value if taken individually.

In summary, we have revealed several associations between EGFR pathway related markers and shown that combining various EGFR pathway related markers into a prognostic model improves prediction relative to the prognostic model based on the pTNM staging system alone. Although the improvement in prognostic value of the markers over pTNM stage is modest, our results suggest that a combined immunohistochemical analysis of PTEN, pSTAT5, HIF1α and pCMET (residue 1349) could be used to refine prognosis in NSCLC patients. These markers should be therefore carefully considered in studies where treatment with EGFR inhibitors are given, in the adjuvant setting in resected patients and in advanced disease.
REFERENCES


[17] Lu Y, Liang K, Li X, Fan Z. Responses of cancer cells with wild-type or tyrosine kinase domain-mutated epidermal growth factor receptor (EGFR) to EGFR-targeted therapy are linked to downregulation of hypoxia-inducible factor-1 alpha. Molecular Cancer 2007; 6(1): 63.


INTEGRATION OF GENE DOSAGE AND GENE EXPRESSION IN NON SMALL CELL LUNG CANCER; IDENTIFICATION OF HSP90 AS POTENTIAL TARGET

Mariëlle I Gallegos Ruiz, Karijn Floor, Paul Roepman, José A. Rodriguez, Gerrit A. Meijer, Wolter J Mooi, Ewa Jassem, Jacek Niklinski, Thomas Muley, Nico van Zandwijk, Egbert F Smit, Kristin Beebe, Len Neckers, Bauke Ylstra and Giuseppe Giaccone

PLoS ONE 2008
In press
IDENTIFICATION OF HSP90 AS POTENTIAL TARGET IN NSCLC

ABSTRACT

Background: Lung cancer causes approximately 1.2 million deaths per year worldwide, and non-small cell lung cancer (NSCLC) represents 85% of all lung cancers. Understanding the molecular events in non-small cell lung cancer (NSCLC) is essential to improve early diagnosis and treatment for this disease.

Methodology and principal findings: In an attempt to identify novel NSCLC related genes, we performed a genome-wide screening of chromosomal copy number changes affecting gene expression using microarray based comparative genomic hybridization and gene expression arrays on 32 radically resected tumor samples from stage I and II NSCLC patients. An integrative analysis tool was applied to determine whether chromosomal copy number affects gene expression. We identified a deletion on 14q32.2-33 as a common alteration in NSCLC (44%), which significantly influenced gene expression for HSP90, residing on 14q32. This deletion was correlated with better overall survival (P=0.008), survival was also longer in patients whose tumors had low expression levels of HSP90. We extended the analysis to three independent validation sets of NSCLC patients, and confirmed low HSP90 expression to be related with longer overall survival (P=0.003, P=0.07 and P=0.04). Furthermore, in vitro treatment with an HSP90 inhibitor had potent antiproliferative activity in NSCLC cell lines.

Conclusions: We suggest that targeting HSP90 will have clinical impact for NSCLC patients.
INTRODUCTION

Lung cancer is the leading cause of cancer deaths worldwide [1], and non-small cell lung cancer (NSCLC) represents 85% of lung cancers. A better understanding of the molecular events underlying the development and progression of the disease may contribute to improve clinical management of NSCLC patients. A number of genes, e.g. P53, RAS, P16 and EGFR, have been shown to be altered in NSCLC [2]. Given the heterogeneous and complex nature of this tumor type, it is likely that many genes driving NSCLC tumorigenesis have yet to be identified.

Chromosomal aberrations are thought to be critical events in human tumorigenesis, and several genomic regions frequently harboring DNA gains (3q, 5p, 7q, 8q, 11q and 16p) and losses (3p, 4q, 5q, 6q, 8p 9p and 13q, 17q) have been identified in NSCLC patients [3]. Using array based comparative genomic hybridization (aCGH) and gene expression microarrays, DNA copy number changes and gene expression can be measured throughout the whole genome of tumor cells. By combining the data from these analyses, it is possible to obtain an integrated genome wide view of gene dosage aberrations and their effect on gene expression, which might help in identifying genes important in NSCLC [4].

In the present study, we have performed an integrative analysis of chromosomal copy number and gene expression on radically resected tumor samples from 32 NSCLC patients. Two new algorithms, ’CGH call’[5] and ’ACE-it’[6], were applied to analyze the data. We identified a deletion on chromosome region 14q32.2-33 in 44% of NSCLC patients. This deletion was related with improved patient survival, and was associated with decreased expression of HSP90, a molecular chaperone for several oncoproteins that is being explored as a novel target in anticancer therapy. Low HSP90 expression was correlated with improved survival in the 32 NSCLC patients analyzed initially. Further analysis of three independent sets of NSCLC patients confirmed a significant association between patient survival and HSP90 expression. In addition, in vitro experiments show NSCLC cell lines to be extremely sensitive to the HSP90 inhibitor 17-AAG. Our data suggest an important role for HSP90 in NSCLC.

METHODS

Patients and samples

The test set consisted of radically resected tumor specimens of 32 early stage NSCLC patients. Three patients had a survival time of less than 30 days and were considered postoperative deaths. Therefore these three patients are not included in the survival analyses. Patients had a median follow up of 86 months (range 0.4-135.5). Verbal informed consent had been obtained from all patients and handling of samples was in accordance with protocols approved by the
IDENTIFICATION OF HSP90 AS POTENTIAL TARGET IN NSCLC

The first validation set consisted of 140 radically resected NSCLC patients from the European lung cancer consortium. Patients had a median follow up of 35 months. All patients included had had no prior malignancy, pathological tumor stage 1 or 2 (T1-2), node stage 0+1 (N0-1), no distant metastasis (M0) at time of operation, and no residual disease after resection (R0). None of these patients received (neo)adjuvant chemo- or radiotherapy. The second validation consisted of 111 early stage NSCLC patients from Bild et al. [7]. The third validation set consisted of the publicly available “datasets 1 and 2” from Lu et al. [8] and contained 54 early stage NSCLC patients. A full description of patient characteristics of all four patient sets is provided in Table 1.

### Table 1. Clinical characteristics of test and validation patient sets

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Test set n=32</th>
<th>Validation set 1 n=140</th>
<th>Validation set 2 n=111</th>
<th>Validation set 3 n=54</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24 (75)</td>
<td>105 (75)</td>
<td>24 (44)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8 (25)</td>
<td>35 (25)</td>
<td>29 (54)</td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>13 (41)</td>
<td>43 (31)</td>
<td>58 (52)</td>
<td>14 (26)</td>
</tr>
<tr>
<td>Squamous Cell Carcinoma</td>
<td>15 (47)</td>
<td>78 (56)</td>
<td>53 (48)</td>
<td>36 (67)</td>
</tr>
<tr>
<td>Large Cell Carcinoma</td>
<td>3 (9)</td>
<td>7 (5)</td>
<td>0 (0)</td>
<td>4 (7)</td>
</tr>
<tr>
<td>Others</td>
<td>1 (3)</td>
<td>12 (8)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>0 (0)</td>
<td>7 (5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Former</td>
<td>16 (50)</td>
<td>57 (41)</td>
<td>27 (24)</td>
<td>7 (13)</td>
</tr>
<tr>
<td>Current</td>
<td>11 (34)</td>
<td>65 (46)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5 (16)</td>
<td>11 (8)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Tumor stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>14 (44)</td>
<td>25 (18)</td>
<td>30 (27)</td>
<td>47 (87)</td>
</tr>
<tr>
<td>IB</td>
<td>9 (28)</td>
<td>68 (48)</td>
<td>27 (24)</td>
<td>7 (13)</td>
</tr>
<tr>
<td>IIA</td>
<td>2 (6)</td>
<td>5 (4)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>IIB</td>
<td>7 (22)</td>
<td>42 (30)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Age at diagnosis - years</td>
<td>67 (22-78)</td>
<td>64 (37-79)</td>
<td>NA</td>
<td>66 (48-81)</td>
</tr>
<tr>
<td>Overall Survival - months</td>
<td>34 (0.4-125.5)</td>
<td>35 (0.5-156)</td>
<td>31 (1-87.5)</td>
<td>50 (2-81)</td>
</tr>
</tbody>
</table>

Percentages that do not reach 100% indicate missing data; NA= information not publicly available.
Isolation of genomic DNA and array Comparative Genomic Hybridization

Cryo-sections of frozen tissue samples, flanking the sections used for RNA and DNA isolation, were verified by the study pathologist (WM) to contain at least 50% of tumor cells. Genomic DNA was extracted from each sample using Trizol following manufacturer instructions (Life Technologies, Breda, The Netherlands). DNA labeling and hybridization on CGH 30K oligonucleotide microarrays was performed as described by van den IJssel et al [9].

RNA isolation and gene expression micro arrays

RNA isolation and cDNA labeling followed standard protocols. Hybridization was performed on Agilent platform according to standard procedures described by the manufacturer and elsewhere [10].

Data analysis

For array CGH, spot analysis and quality control were performed using BlueFuse version 3.2 (BlueGenome, Cambridge, UK). Breakpoints, gains, losses and amplifications were detected using the algorithm CGH call. This algorithm converts raw log2ratios to absolute measures of "loss", "normal", "gain" or "amplification" by applying a segmentation algorithm combined with a probability mixture model [11]. In order to statistically test whether gene expression was affected by gene dosage we applied an array CGH expression integration tool, ACE-it[12], in which the called array CGH and normalized log10 ratios for expression arrays were used as input data. ACE-it uses the one-sided Wilcoxon rank sum statistics to test which chromosomal copy number aberrations recurrently affect RNA expression. Calculated p-values are adjusted for multiple testing using the Benjamini Hochberg method [13]. ACE-it only tests genes that meet the criteria of contamination and balance, which are controlled through a threshold on the number of samples. Here the threshold was set at a fixed default setting of 9 samples, meaning that only those chromosomal positions were taken into account that had at least 9 samples in one CGH calling status and no more than 9 in the other status. The entire array CGH data set of the test series (n=32) is available at the GEO database (http://www.ncbi.nlm.nih.gov/projects/geo, accession number GSE7878). The gene expression data of both the test set (n=32) and validation set 1 (n=140) for the 359 genes identified by ACE-it is available at the Array Express database (www.ebi.ac.uk/areep/login, accession number Array design: A-MEXP-749, Experiment data: E-TABM-270). Gene expression data of validation set 2 was obtained using Affymetrix Hu133plus2 chips (GEO accession number GSE3141). The mean value of the MAS5 calculated signal intensities of four probesets detecting HSP90AA1 was used in our calculations.
The third validation set contained gene expression data from Affymetrix Hu95 and Hu133 chips (GEO accession number GSE6253). The Hu95 chip contained one probeset detecting HSP90AA1 and the Hu133 chip contained four probesets, of which the mean value of the RMA calculated signal intensities was used in our calculations.

**Statistics**

An univariate cox regression analysis was performed to investigate relation of gene expression values with survival time. Survival curves were constructed using the Kaplan Meier method and differences in overall were evaluated using the log-rank test. In the test set, three patients with less than 30 days survival time were excluded from the survival analysis, as their death was considered surgical mortality. To determine the independent effects of HSP90 expression, histologic subtype, tumor stage, age and gender, a multivariate cox regression analysis was performed. A P value of less than 0.05 was considered statistically significant.

**Multiplex Ligation dependent Probe Amplification (MLPA)**

For Multiplex Ligation dependent Probe Amplification (MLPA), the subtelomere probe set P070 (MRC Holland, Amsterdam, The Netherlands) containing a probe located within the band 14q32.33 (region 104874216 to 105070384) was used. MLPA was performed according to manufacturer’s instructions using 100ng DNA as input. DNA isolated from blood of a pool of healthy donors was used as reference sample. Probe signals were normalized by dividing the peak area of chromosome 14 by the peak area of chromosome 14p. MLPA generated 14q/14p ratios are plotted against the mean normalized log2ratios of the oligos in area 104787271-105071522 from the array (total of 11 oligos). This region covers the region of the MLPA probe.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>Band</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>End</td>
</tr>
<tr>
<td>3</td>
<td>55285625</td>
<td>124794922</td>
</tr>
<tr>
<td>9</td>
<td>190500861</td>
<td>244968392</td>
</tr>
<tr>
<td>10</td>
<td>237445</td>
<td>50384805</td>
</tr>
<tr>
<td>7</td>
<td>8163179</td>
<td>209312147</td>
</tr>
<tr>
<td>10</td>
<td>102229012</td>
<td>123576527</td>
</tr>
<tr>
<td>12</td>
<td>129933159</td>
<td>137018487</td>
</tr>
<tr>
<td>8</td>
<td>56600324</td>
<td>134194895</td>
</tr>
<tr>
<td>4</td>
<td>201402</td>
<td>8164020</td>
</tr>
</tbody>
</table>

**Table 2. Chromosomal regions with gains and losses present in >20% of NSCLC patients analyzed**

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>Band</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>End</td>
</tr>
<tr>
<td>1</td>
<td>794287</td>
<td>6218763</td>
</tr>
<tr>
<td>2</td>
<td>218493156</td>
<td>220321649</td>
</tr>
<tr>
<td>3</td>
<td>317583737</td>
<td>317673711</td>
</tr>
<tr>
<td>4</td>
<td>24976</td>
<td>2735138</td>
</tr>
<tr>
<td>5</td>
<td>34521369</td>
<td>44910245</td>
</tr>
<tr>
<td>6</td>
<td>142432747</td>
<td>141985957</td>
</tr>
<tr>
<td>7</td>
<td>127225225</td>
<td>13025821</td>
</tr>
<tr>
<td>8</td>
<td>123526352</td>
<td>13025821</td>
</tr>
<tr>
<td>9</td>
<td>132582632</td>
<td>130257936</td>
</tr>
<tr>
<td>10</td>
<td>189965</td>
<td>3164142</td>
</tr>
<tr>
<td>11</td>
<td>650502841</td>
<td>87592966</td>
</tr>
<tr>
<td>12</td>
<td>131534536</td>
<td>131915096</td>
</tr>
<tr>
<td>14</td>
<td>89168871</td>
<td>100051451</td>
</tr>
<tr>
<td>15</td>
<td>37491</td>
<td>4794626</td>
</tr>
<tr>
<td>16</td>
<td>27639900</td>
<td>31028401</td>
</tr>
<tr>
<td>17</td>
<td>66038796</td>
<td>6659430</td>
</tr>
<tr>
<td>18</td>
<td>82507653</td>
<td>88176752</td>
</tr>
<tr>
<td>19</td>
<td>33712</td>
<td>1512609</td>
</tr>
<tr>
<td>20</td>
<td>72626307</td>
<td>78046922</td>
</tr>
<tr>
<td>21</td>
<td>395096</td>
<td>13451422</td>
</tr>
<tr>
<td>22</td>
<td>54738029</td>
<td>651747941</td>
</tr>
<tr>
<td>23</td>
<td>60513620</td>
<td>60110730</td>
</tr>
<tr>
<td>24</td>
<td>60513599</td>
<td>63746634</td>
</tr>
<tr>
<td>25</td>
<td>3158113</td>
<td>3846872</td>
</tr>
<tr>
<td>26</td>
<td>58760290</td>
<td>62397233</td>
</tr>
<tr>
<td>27</td>
<td>42368792</td>
<td>46688780</td>
</tr>
<tr>
<td>28</td>
<td>36296652</td>
<td>30571684</td>
</tr>
</tbody>
</table>
Quantitative RT-PCR

In order to validate expression values obtained via gene expression arrays, we performed quantitative real-time PCR using Taqman® technology and the ABI PRISM 7500 Sequence Detection System instrument equipped with the SDS version 1.3.0 software (Applied Biosystems, Foster City, CA, USA). Forward and reverse primers and probes were designed and produced by Applied Biosystems for HSP90AA1 (Hs00743767_sH), and for the endogenous control gene GUSB (Hs00939626_mi). PCR was carried out in a 25-μl reaction volume that contained 50 ng of cDNA, 1x TaqMan Universal PCR Master Mix, and the primer and probe sets for HSP90AA1 and GUSB. Each sample was analyzed in duplicate, and the average threshold cycle (Ct) values of each sample for GUSB, were subtracted from the average Ct values for HSP90AA1. Taqman-generated ΔCt values were log transformed to expression values and plotted against the Δlog2ratios between GUSB and HSP90AA1 from the expression array.

Cell growth inhibition studies

Growth inhibition following in vitro exposure to the clinically used Hsp90 inhibitor 17-AAG (InvivoGen, San Diego, CA) was examined in 4 EGFR wild type NSCLC cell lines (H460, H157, H441 and A549, obtained from Drs. P. Dennis and F. Kaye, NCI, Bethesda, MD). Briefly, 10^5 cells were seeded in 6-well tissue culture plates (Sigma-Aldrich, St. Louis, MO), allowed to adhere, and then continuously exposed to various concentrations of 17-AAG (0, 10, 30, 100, 300, 1000 nM) for 24, 48, or 72 hours. At each time point, cells were detached from the wells, incubated with trypan blue, and viable (trypan blue-excluding) cell number was determined in triplicate using a hemacytometer. The mean cell number with standard error bars is shown at each time point. In addition, the IC_{50} (drug concentration at which 50% growth inhibition is obtained) of 17-AAG at 72 hours was determined for each cell line.

RESULTS

Gene dosage-related gene expression changes in NSCLC

Chromosomal aberrations were abundant in the 32 NSCLC patients analyzed. In order to identify breakpoints of gains and losses we applied the algorithm CGH call [14]. In Table 2 the chromosomal regions in which gains or losses were present in at least 20% of patients are listed. The statistical tool ACE-it [15] was used to determine whether gene copy number affected gene expression. A total of 359 transcripts turned out to be significantly affected by copy number. In Figure 1 the areas of affected genes are indicated for 32 NSCLC patients, shown in green (gained regions) or red (lost regions).
An univariate cox regression survival analysis was performed for expression of all 359 genes that were identified to be influenced by copy number (see supplementary table S1). After multiple testing correction (using the Benjamini Hochberg method) none of 359 genes remained significant for survival on this small patient set (n=32). The top list of genes correlated with survival (ranked on raw p-values) contained mainly genes located on chromosomes 3 and 5 gained regions. We also observed one gene in the top-20 list, HSP90AA1, located on chromosome 14. The HSP90AA1 gene (generally referred to as HSP90), located on 14q32.2, was the only gene in this region with significantly reduced expression in patients affected by loss of this region (P=0.05). These observations prompted us to investigate this locus in more detail.

**Genomic aberrations on 14q and HSP90AA1 gene expression correlation with survival**

We investigated the correlation of the recurrent tight deletion at region 14q32.2-33, with survival. Interestingly, patients in whom this region was deleted had an improved overall survival (OS) compared to patients harboring a normal gene dosage at this locus (5 year OS 69% vs. 41%, P=0.004 - Figure 2A).

In order to investigate HSP90 expression in relation with survival, we divided the 32 patients into two groups based on their survival status two years after tumor resection. About half of the patients were categorized as "high-risk" and half as "low-risk". To identify whether both risk
groups could be discriminated using gene expression of HSP90, we constructed Kaplan Meier survival estimates for two equally large groups with “low” and “normal”, based on the median HSP90 expression. Low expression of HSP90 was associated with better overall survival (5 year OS 70% vs. 40%, P= 0.13 – Figure 2B) although differences between the two groups were initially not significant due probably to a combination of a low magnitude of the difference and a low number of patients analyzed.

**FIGURE 2.** Loss of 14q32.2-32.33 chromosomal region and HSP90 expression in relation with survival. (A) Kaplan-Meier curves for overall survival are shown for 29 patients in relation to gene dosage in chromosome region 14q32.33. (B) Overall survival for 29 patients in relation to HSP90 expression. Low expression was defined as expression lower than the median of the total 32 samples, “normal” expression was defined as higher than the median of 32 samples. Three of 32 patients included in the analysis of gene dosage and expression were excluded from the survival analysis because of a survival time of less than 30 days.

**FIGURE 3.** HSP90 expression and survival in three validation sets of NSCLC patients. Overall survival for (A) 140 patients with NSCLC, validation set 1 (B) 111 NSCLC patients, validation set 2 and (C) 54 patients with NSCLC, validation set 3. The cut off for distinction between low and “normal” expression was based on the 33-percentile of expression values.
Technical validation of 14q deletion and HSP90 expression

To validate the deletion observed in region 14q32.2-33 we performed multiplex ligation dependent probe amplification [16] (MLPA) analysis using a subtelomere probeset containing a probe in region 14q32.33. The correlation coefficient ($R^2$) between the loss of this area detected by array CGH and by MLPA was 0.439.

To validate the HSP90 expression data obtained with microarrays we performed quantitative RT PCR using the Taqman® technology. A good correlation between the expression of HSP90 measured with the two different techniques was observed ($R^2=0.6431$).

Validation of HSP90 expression and relation with survival in independent patient series

To further investigate the association between low HSP90 expression and NSCLC patient prognosis, we used three independent validation sets of NSCLC patients. In all three patient sets, the “low-risk” / “high-risk” patient distribution across patient cohorts was approximately two-thirds vs. one-third. Therefore, we used the 33-percentile of HSP90 expression as cut off for separation of patients with “normal” (i.e. high-risk) and “low” (i.e. low-risk) expression. For all three validation sets, low expression of HSP90 was correlated with improved overall survival. This correlation was significant for the first and third validation sets ($P=0.003$ and $P=0.04$), and borderline significant for the second set ($P=0.07$) (Figure 3). Multivariate analysis revealed that HSP90 prognostic value was independent from the stage, histologic subtype, age and gender of patients (Table 3).

<table>
<thead>
<tr>
<th>Table 3. Multivariate cox analysis of HSP90 expression and survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation set 1</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>HSP90</td>
</tr>
<tr>
<td>95% CI</td>
</tr>
<tr>
<td>P value</td>
</tr>
<tr>
<td>STAGE</td>
</tr>
<tr>
<td>95% CI</td>
</tr>
<tr>
<td>P value</td>
</tr>
<tr>
<td>HISTOLOGY</td>
</tr>
<tr>
<td>95% CI</td>
</tr>
<tr>
<td>P value</td>
</tr>
<tr>
<td>AGE</td>
</tr>
<tr>
<td>95% CI</td>
</tr>
<tr>
<td>P value</td>
</tr>
<tr>
<td>GENDER</td>
</tr>
<tr>
<td>95% CI</td>
</tr>
<tr>
<td>P value</td>
</tr>
</tbody>
</table>

HR=Hazard Ratio; n/a=data not publicly available or incomplete
Validation of Hsp90 as a viable molecular target in a panel of NSCLC cell lines

Hsp90 is a molecular chaperone that stabilizes several oncoproteins, including EGFR, and constitutes a novel potential target for anticancer therapy. The data above showed that Hsp90 expression level is a prognostic indicator of long-term survival in a large series of NSCLC patients, and suggest that Hsp90 inhibitors may have broader utility in this disease than previously recognized. To examine this possibility, we tested whether pharmacologic inhibition of Hsp90 function would impact the in vitro cell growth of a panel of NSCLC cell lines. The growth of these cell lines, all bearing wild-type EGFR, was profoundly inhibited, in a dose- and time-dependent manner, by sub-micromolar concentrations of 17-AAG, an Hsp90 inhibitor [17] currently in phase II clinical trial (Figure 4). At 72 hours, the IC50 was below 50 nM 17-AAG in all cases, and higher drug concentrations consistently resulted in marked cytotoxicity.

**FIGURE 4. Sensitivity of a panel of wild type EGFR-expressing NSCLC cell lines to the Hsp90 inhibitor 17-AAG.** (A-D) time- and dose-dependent inhibition of the in vitro growth of H460, H157, H441, and A549 NSCLC cell lines following exposure to 17-AAG. Cells were seeded at 10^5/well, and viable cell number was determined on subsequent days as described in Methods. 17-AAG concentrations at (H441) or above (A549, H460 & H157) 30 nM uniformly resulted in time-dependent loss of cell viability. The IC50 value of 17-AAG (continuous exposure for 72 h) for each cell line is as follows: H460 = 30 nM, H157 = 15 nM, H441 = 8 nM, and A549 = 20 nM. A colored version of this figure can be found in appendix C – page 208.
DISCUSSION

Based on array-CGH data in NSCLC, it has been shown that multiple molecular carcinogenesis pathways exist that are most likely related to gender and smoking habits [18]. Furthermore, it was shown that there is a large overlap between aberrations observed in the adenocarcinoma and squamous cell carcinoma subtype, except for 3q gains which seem to be more specific for the squamous cell carcinoma subtype [19]. Various gene expression signatures have been correlated to survival of NSCLC patients [20-22]. In addition, molecular studies have allowed the development of personalized treatment approaches in several tumor types [23-25]. However, it is likely that many cancer-related target genes have not been identified yet. In this regard, integrated genome wide screening of copy number changes and gene expression using microarrays has been recently carried out in various tumor types to identify genes whose expression is affected by gene dosage [26-30]. These studies aim to identify novel cancer-related genes and to define novel biomarkers for response or prognostic signatures. In both NSCLC and ductal pancreatic cancer, two focal amplifications of 8p12 and 20q11 have been studied in detail leading to two candidate genes (WHSC1L1 and TPX2) important in these diseases [31].

We here performed an integrated genome wide screening of gene copy number changes and gene expression in 32 radically resected NSCLC patients, in order to identify novel NSCLC-related genes. By using ACE-it, a novel informatics tool for integration of gene dosage and gene expression data, we identified 359 transcripts to be significantly affected by copy number. A cox survival analysis on all 359 genes revealed no significant relation after multiple testing. The top list of genes related to survival mainly included genes residing on gained regions 3q and 5p. These regions cover many genes and to pinpoint the gene of most importance is a challenging task. Further investigations should elucidate the importance of these genes in relation to NSCLC, in particular of those in the top list of correlation with survival such as SLC45A2, WDR70 and NIPBL (see supplementary table S1). In this paper we focused on the recurrent deletion on chromosome 14 and the gene HSP90, which was also in the top list of relation with survival and not previously investigated in detail. Deletion of region 14q32.2-33 was correlated with improved survival, further suggesting that it may contain one or more genes related to NSCLC progression. Deletion of this region has been previously described by one group reporting genomic aberrations in NSCLC, but was not investigated in further detail [32]. Out of the 109 genes mapping to the 14q32.2-33 region, HSP90 was the only gene with significantly lower expression in patients harboring the 14q32.2-33 deletion. In the initial series of 29 patients (3 patients excluded from survival analysis), we observed improved survival in patients with lower levels of HSP90. The association between HSP90 expression levels and NSCLC patient prognosis was confirmed to be
significant in three independent validation sets of NSCLC patients. Multivariate analysis including stage, histology, age and gender showed that HSP90 remained independently related to survival. A critical issue in defining "low" and "normal" expression is the choice of an appropriate cut off value. In the initial analyses we used the median of expression ratios as cut off between "low" and "normal" expression, since the low-risk and high-risk separation of patients was equal. However, in the validation sets the low-risk and high-risk survival groups were not equally balanced (two thirds versus one third). Consequently the cut off values used in these data sets was not the median value, but the 33-percentile.

In this study, using a genome wide integrative analysis of gene copy number and expression we were able to identify expression of HSP90 as an important gene in early stage NSCLC patients. HSP90 is a chaperone protein involved in the stabilization of multiple oncoproteins such as EGFR, Her-2 and Akt [33]. Recent work has shown that HSP90 plays a role in maintaining the active conformation of EGFR and in particular EGFR mutants [34,35]. We show here that several NSCLC cell lines bearing wild-type EGFR are sensitive to HSP90 inhibition, indicating that the inhibitory effect of 17-AAG can not be solely attributed to mutant EGFR. HSP90 has been recently recognized as a potential cancer therapeutic target and investigations of HSP90 inhibitors are ongoing [36,37]. In this regard, glioblastoma cells overexpressing EGFR, but resistant to inhibition by EGFR kinase inhibitors, were sensitive to HSP90 inhibition [38]. Therefore, while Hsp90 may be required to stabilize over-expressed or mutated EGFR, our data support a more wide-ranging and complex role for Hsp90 in mediating NSCLC growth and survival. The low nanomolar sensitivity observed in the 4 cell lines tested in our experiments, is in agreement with other published reports of highly 17-AAG sensitive tumor cell lines [39-41]. These concentrations are readily achievable in patients for prolonged time periods using current scheduling and dosing regimens [42].

In summary, the observation that HSP90 expression level is a prognostic factor for NSCLC patient survival (independent of EGFR mutational status), coupled with the extreme sensitivity of EGFR wild type NSCLC cells to the Hsp90 inhibitor 17-AAG, suggests that Hsp90 inhibitors may have greater clinical utility in NSCLC than has been previously considered and warrants further investigation of the dependence of other proto-oncogenes on this chaperone protein in NSCLC.
REFERENCES


ACKNOWLEDGEMENTS

We thank the European lung cancer consortium for providing microarray gene expression data of the 140 independent NSCLC patients; Netherlands Cancer Institute: Sjaak Burgers, Tony van de Velde; Medical University of Gdansk: Jacek Jassem, Amelia Szymanowska, Marcin Skrzypski, Barbara Szostakiewicz, Witold Rzyman; Medical University of Białystok: Jerzy Laudanski, Mirosław Kozłowski, Lech Chyczewski; Thoraxklinik Heidelberg: Michael Meister, Philipp A. Schnabel, Henrik Dienemann and Hans Hoffmann. We also give our special thanks to Sjoerd Vosse (Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands) for helpful bioinformatics discussions and to Hans Gille (Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands) for performing MLPA analysis.
CHAPTER 10

SUMMARIZING DISCUSSION

&

FUTURE PERSPECTIVES
SUMMARIZING DISCUSSION & FUTURE PERSPECTIVES

Most common cancers have a large number of genetic alterations, some of which are responsible for their invasive and metastatic phenotype. Genetic alterations can be mutations, chromosomal rearrangements, DNA copy number abnormalities or epigenetic changes, such as methylation or micro RNA regulation. These aberrations can all have an effect on gene expression, and eventually on the expression or function of proteins. This thesis aimed to investigate specific genetic aberrations in patients with Non-Small Cell Lung Cancer (NSCLC). The research described in this thesis is focused on the identification and use of such aberrations as ‘biomarkers’ to predict response to therapy or prognosis, and to identify candidate genes for targeted therapy in NSCLC patients.

The epidermal growth factor receptor (EGFR) is a gene in which mutations and high DNA copy number have been described to be correlated with response to EGFR tyrosine kinase inhibitors (TKIs) [1-3]. In contrast, mutations in the KRAS oncogene have been related with primary resistance to these agents. Secondary resistance to EGFR TKIs has been attributed, at least partially, to a specific T790M EGFR mutation [4]. We have confirmed the correlation between the presence of EGFR and KRAS mutations and correlation with response to Erlotinib (an EGFR TKI) in a clinical phase II study, described in chapter 2. However, in our study as in the literature, the concordance between response and presence of EGFR and KRAS mutations was not perfect [5-7]. The lack of a perfect concordance is further emphasized by a case report that is described in chapter 3. This case report describes a Caucasian, male, former smoker, with NSCLC and squamous cell carcinoma histology, responding to Erlotinib. The clinical characteristics formerly reported to be associated with positive response to EGFR TKIs are female gender, Asian ethnicity, never-smoking status and adenocarcinoma histology [8,9]. As chapter 3 describes there are always exceptions to the rule, highlighting that care should be taken to base treatment with EGFR TKIs on those clinical characteristics only. Recent data [10] cast doubts that these clinical characteristics and EGFR aberrations are specific to EGFR TKI benefit and may be a more general indication of benefit to other treatments as well.

EGFR and KRAS mutation analyses can also be applied in a different clinical situation. Some patients with NSCLC present with multiple lung lesions at time of diagnosis or develop additional lesions during their lifetime. To decide on treatment, it is important to determine the relation between the various lesions, in order to discriminate between second primary tumors or metastasis. Currently, this is done by pathologic examination of histology slides, although this is sometimes inconclusive. A second case report (chapter 4), describes three patients who developed multiple lung lesions. We performed EGFR and KRAS mutation analysis of their
different tumors and, in addition, performed array CGH on the multiple lesions. By combining those techniques we were able to define the clonal relation between the multiple lesions.

Various techniques to determine EGFR mutations have been investigated, although the most commonly used method remains PCR and sequencing [11-15]. The source of the samples, i.e., whether the DNA or RNA was isolated from formalin fixed and paraffin embedded (FFPE) or fresh frozen samples, has been shown to influence outcome of mutation analysis [16]. We have performed a comparison study (chapter 5), in order to determine the effect of using different source, and of using DNA or RNA in the mutation analysis using PCR and sequencing. We showed that RNA extracted from frozen tissue is the preferred source for analysis since all PCRs were successful and, in some cases, EGFR mutations could be more easily identified using RNA compared to DNA. The latter could not have been caused by the percentage of tumor cells present in the sample, since RNA and DNA were derived from the same tissue. It has been reported that in EGFR amplified tumors there is selective amplification of the mutant allele [17]. The fact that we have observed the EGFR mutation more prominently in the RNA sample suggests that the mutant allele may be transcribed more efficiently. Chapter 5 furthermore highlights the challenge to standardize fixation and storage protocols in order to allow molecular studies to be completed without sample related technical difficulties.

As mentioned previously, high EGFR copy number can also be a predictive marker for response of EGFR TKIs [18,19]. The technique used to assess copy number in most studies is fluorescent in situ hybridization (FISH). This technique has some technical limitations since the use of fluorescence requires the need for an expensive fluorescence microscope, and rapid evaluation because of fading of the signal. These limitations can be overcome by a method that uses a chromogenic rather than a fluorescent signal, chromogenic in situ hybridization (CISH). We performed a comparison study, described in chapter 6, in order to evaluate whether CISH could be a useful strategy to determine EGFR copy number in NSCLC samples. We determined the concordance between FISH and CISH, and evaluated the performance of the two techniques in both FFPE and in fresh frozen samples. We show that CISH is a useful strategy to evaluate EGFR copy number and showed high concordance with FISH. This chapter also highlights the challenge to standardize interpretation of in situ hybridization techniques and the urge to determine the appropriate cut-off in defining to what extent EGFR copy number influences response to anti-EGFR TKI treatment.

Various types of EGFR mutations have been identified to date [20], and they are differentially related to response to EGFR TKI treatment. In this regard, it has been shown that deletions in EGFR exon 19 are related with better response to EGFR TKIs than the L858R mutation [21]. Moreover, the T790M mutation is related with acquired resistance to EGFR TKIs [22,23]. Yet another type of EGFR mutations, insertions in EGFR exon 20, have been related to primary
resistance to EGFR TKIs [24]. These data highlight that treatment cannot be simply based on the
presence of any type of EGFR mutation, but each mutation should be investigated in more detail.
To circumvent the use of complex and time consuming assays [25,26], we describe in chapter 7
a simple assay able to quickly determine the influence of specific mutations on sensitivity to EGFR
TKIs. Using mutagenesis of a YFP-tagged fragment of the EGFR intracellular domain, following
transfection and immunofluorescence microscopy analysis we were able to determine the effects
of Erlotinib treatment. Using this system, we confirmed the sensitivity to Erlotinib of the EGFR
deletion mutant and the L858R mutant, and the resistance to Erlotinib of the T790M mutant.
Moreover, we show the insensitivity to Erlo tin ib of some uncommon EGFR mutations that were
identified during a mutational screening of NSCLC patients. This system might be important in
the clinical management of NSCLC patients that present with uncommon EGFR mutations.
Moreover, this assay can be applied in investigations of genes and drugs, other than EGFR and
EGFR TKIs.
The EGFR pathway has extensively been studied to identify additional biomarkers for response to
EGFR TKIs. In addition to the previously mentioned EGFR and KRAS mutations and EGFR copy
number, studies have evaluated the activation status of AKT and STAT signaling pathways [27],
and the expression level of HER2 [28,29]. More recently, amplification of the MET receptor
[30,31], the expression of epithelial to mesenchymal transition (EMT) markers, such as E-
cadherin and vimentin [32,33], and the downregulation of HIF1α have also been linked to
responsiveness to EGFR targeted agents [34]. Next to their predictive role for response to EGFR
TKIs, some of those markers have also been linked to prognosis. However, conflicting results are
often reported by different laboratories [4,35-42], and there is thus no single marker that can be
used to predict prognosis of NSCLC patients. In chapter 8 we evaluate whether the combined
analysis of several markers provides better prognostic information than the evaluation of single
markers or the pTNM staging system, the most reliable prognostic marker so far, in which the
size of the tumor and the degree of spread to lymph nodes and other sides in the body is taken
into account. We investigated EGFR and KRAS mutational status, EGFR copy number, and the
expression of EGFR, HER2, pCMET, pAKT, PTEN, pSTAT3, pSTAT5, pERK, HIF1α, E-CADHERIN
and VIMENTIN, using immunohistochemistry on tissue microarrays (TMA) in resected NSCLC
patients. Using a statistical model we have compared the performance to predict survival using
the molecular markers with the pTNM staging system. We show that using the molecular markers
we were able to improve prediction of prognosis by 7-10% as compared to the pTNM staging
system alone. Furthermore, this statistical model was able to identify the most important
markers, which were pSTAT5, HIF1α, PTEN and pCMET. Although the improvement in prediction
is only modest, this study highlights that a combined analysis of several molecular markers and
pathological staging can provide better prognostic information than the use of single markers.
Although the EGFR pathway is extensively studied in NSCLC, and well recognized to play an important role in this disease, there is also a subset of patients in which EGFR probably does not play a prominent role. Thus, new genes need to be identified. The use of microarray comparative genomic hybridization (array CGH) has been used to identify that amplification of the gene encoding the MET receptor is partially responsible for developing acquired resistance to EGFR TKIs [30,31]. In chapter 9, we describe a study in which we used array CGH to determine genome-wide copy number changes in NSCLC patients. Using the same patient samples, gene expression profiling was performed. Through the combined analysis of the results obtained with both techniques, and through the use of newly developed algorithms to analyze the data [43,44], we were able to identify genes whose expression was influenced by copy number. We identified 359 genes, mainly residing on chromosomes 3 and 5, showing altered expression as a result of aberrant gene copy number. In this chapter we focus on the gene HSP90 residing on chromosome 14, whose expression showed a clear relation with the survival of patients. This gene was the only gene whose expression was affected by a deletion of this tight region. This study further stresses the importance that specific inhibitors of HSP90, which are already in clinical investigation [45,46], may have in the clinical management of NSCLC patients. As described in chapter 9, microarray studies to determine gene expression and gene copy number can be used to identify potential targets for therapeutic purposes. Other use of microarray analysis in cancer research is to compare patient groups with known outcome (treatment, relapse, survival) and to determine a signature able to discriminate between groups of patients (supervised analysis). One can also use the gene expression or array CGH data itself to create groups, which could be related with patient outcome (unsupervised analysis). The supervised approach was successfully used in studies that identified gene signatures related with lung cancer patient outcome. An interesting observation was the fact that there were only some genes in these signatures that overlapped [47-52]. A similar phenomenon is observed in breast cancer patients [53,54]. The genes included in the lung cancer signatures, as well as in the breast cancer signatures, were often part of similar pathways, including pathways regulating immune response, cell growth and DNA repair. This highlights another application of microarrays, namely the identification of important pathways in selected groups of patients. Some important aspects when analyzing microarray data should be considered. Microarray experiments result in a huge amount of information, and care should be taken to analyze the data in the most efficient way without misinterpretation or missing important findings. One important aspect is that the study should be clearly designed and objectives clearly formulated. It is very important to correct for confounding factors such as stage, histologic subtype, age and treatment of the patient. Another important aspect is the use of the appropriate algorithm and, since microarray data will always result in more variables than cases, multiple testing corrections.
should always be applied. After having determined a certain signature or genes of interest, validation of the findings needs to be carried out in an independent set of patients. For example, Wigle et al. reported the use of a gene signature to predict recurrence in NSCLC patients [55], and the same group performed a validation study. Although RT PCR was able to confirm the results in the same set, a microarray analysis of an independent set of patients failed to confirm the prognostic significance of the initial microarray results [56]. Our study, described in chapter 9, of this thesis is also currently being validated in an independent set of NSCLC patients.

A nice example of a well validated study is the 70-gene signature for breast cancer [53], of which results have been confirmed in two independent patient sets [57,58], and an update of the original series was recently published [59]. These studies have led to approval by the US Food and Drug administration of the 70-gene signature (Mammaprint®) in breast cancer patients.

The introduction of the publicly available microarray databases has allowed the elaboration of meta-analysis. One such analysis is the study described by Potti et al. in which they constructed genomic signatures to guide the use of chemotherapeutics [60]. They used 60 cell lines derived from various cancer cell types, and developed gene expression signatures that correlate with the response of these cells to various commonly used cytotoxic agents. Using publicly available data, they derived signatures from microarray profiles of the NCI-60 human cancer cell lines with known in vitro sensitivity or resistance to a particular drug. They used these profiles to predict in vivo chemotherapeutic response to seven different drugs. However, a similar analysis was unsuccessfully attempted by another group, who could not reproduce these results [61]. This finding emphasizes that quality control and continued checking of all the bioinformatics and procedures is also extremely important, and highlights that also results from meta-analysis should be treated critically and with caution.

Although great progress has been achieved with the use of microarrays to determine gene expression profiles and copy number aberrations, these techniques do not identify gene regulation by epigenetic changes such as regulation by non-coding RNAs (i.e. micro RNAs) or additional post translational modifications. In this respect, other emerging research fields are micro RNA expression analysis and proteomics. Micro RNAs (miRNAs) are small non coding RNAs, which are expressed in various organisms, including animals, plants, and viruses. These miRNAs are able to regulate several target mRNAs thereby affecting gene expression [62]. Abnormal expression of miRNAs has been shown in various types of cancer [63,64]. In lung cancer, unique miRNA profiles have been linked to patient outcome [65]. The identification of the role of miRNAs represents an area of active investigation [66]. As proteins are the effectors of the information encoded by genes, technologies that look at proteins in an "omic" fashion are also being developed, and proteomic signatures have been described in lung cancer [67-70]. A strategy based on the combined analysis of proteomics and microarrays, and confirmation with
imunohistochemistry on tissue micro arrays appears promising [71,72]. As described in chapter 8 and 9 of this thesis, we have gathered information on the DNA, RNA and protein level simultaneously in several patients and we aim to perform a combined analysis of these parameters in the near future.

In conclusion, the research described in this thesis shows that progress has been made in identifying candidate therapeutic targets and biomarkers for prediction or prognosis purposes in lung cancer. In NSCLC, specifically, EGFR and KRAS mutations have been shown to correlate with response to EGFR TKIs although exceptions remain. We have identified several genes that might have an interesting role in NSCLC, which has to be investigated in more detail in future research. The emerging fields of microarray technology for DNA, RNA and miRNA screening, as well as proteomics, has already given us many new insights into the complexity of human cancers. We are facing many challenges to improve our understanding of cancer and to ultimately improve therapy for cancer patients.
REFERENCES


[34] Lu Y, Liang K, Li X, Fan Z. Responses of cancer cells with wild-type or tyrosine kinase domain-mutated epidermal growth factor receptor (EGFR) to EGFR-targeted therapy are linked to downregulation of hypoxia-inducible factor 1 alpha. Molecular Cancer 2007; 6(1): 63.


APPENDIX A

SUPPLEMENTARY INFORMATION

CHAPTER 8

COMBINED ASSESSMENT OF EGFR PATHWAY-RELATED MOLECULAR MARKERS AND PROGNOSIS OF NSCLC PATIENTS
DESCRIPTION OF RANDOM FOREST SURVIVAL PREDICTION

Given a training data set of survival/censoring times, censoring status and other variables—the covariates—measured on a sample of patients, the \texttt{rsf} function (the main function of the package) constructs an ensemble of survival function estimates—which we may call the predictor. Given the values of the covariates of a patient not in the training data set, the \texttt{predict.rsf} function associates an ‘appropriate’ survival function (determined by the log-rank splitting rule) of the predictor to them, and this survival function can be used to predict the survival time of the ‘new’ patient.

The predictor is a collection of averages of survival curves, each survival curve being based on a bootstrap sample from the training data set. More specifically, each bootstrap sample is stratified into sub-samples, and a survival curve is computed within each stratum; a collection of such survival curves may be called a bootstrap ensemble.

The stratification underlying a bootstrap ensemble involves a relatively complex procedure based on the covariates; what is important to know is that each stratification (i) is based on a randomly selected set of covariates (from a set of covariates fixed by the user) and (ii) is essentially determined so as to maximize pairwise differences between survival curves from different strata. Given the covariates of a ‘new’ individual, one can, for a given bootstrap ensemble, determine the stratum to which it belongs and hence take the survival curve of that stratum as the individual’s survival curve. A better estimate of the survival curve of a new individual, however, is obtained by averaging the individual’s estimates across bootstrap ensembles, which is what the predictor amounts to.

The quality of \texttt{rsf} predictors is assessed in terms of an error rate which takes values between 0 and 1, takes the value 0 if the prediction is 100% accurate, the value 1 if the prediction is 100% inaccurate, and the value 0.5 if the prediction is as accurate as ‘random guessing’. The error rate is defined as one minus Harrell’s concordance index.

Suppose we are given two patients about whom the survival times and statuses are known; and suppose that we have a survival curve estimate for each of these patients (obtained with a training data set). If the survival curve estimates are any good, they should more or less agree with the actual survival times and statuses of the patients. For instance, the fact that the survival curve estimate of patient 1 is ‘more optimistic’ than that of patient 2 should usually be accompanied by the observation that patient 1 actually lived longer than patient 2, though of course the concordance between estimate and observation cannot be expected to hold for all pairs of patients. The percentage of times it does hold, as one goes through all possible pairs of patients, is essentially what we call Harrell’s concordance index.
With the selected patient group we have used ‘holdout method with random subsampling’ to assess the predictive value of some of the protein, genotypic and clinical variables. After randomly selecting two thirds of the data and using them to construct a random survival forest predictor, we computed the error rate of the predictor on the last third of the data. We have repeated this procedure 1000 times and then averaged the error rate to get an estimate of the expected error rate incurred by using a random survival forest predictor.

The other settings (not mentioned in the article) used in the construction of each predictor were:

```
```

Below is an example of the R code used in our calculations.

```
#1. Loading and seeing the data
SurvivalProteins<-read.table("CategorizedData4.txt",header=TRUE,sep=" ")
edit(SurvivalProteins) #Shows the data in the Data Editor

#2. Constructing the predictor with 2/3 of the data, using it to predict the remaining 1/3, and assessing the performance of this procedure
set.seed(2000,kind=NULL)
error.rate<-rep(0,nrep)
threshold<-1 #Percentage of the 'important' variables to be reported in the output
no.patients<-length(SurvivalProteins[,1])
for (i in 1:nrep)
{
p<-2/3 #Proportion to be used as training data set
randomization.vector<-c(rep(1,round(p*no.patients)),rep(0,no.patients-round(p*no.patients)))
randomization.vector<-sample(randomization.vector)
training.sample<-SurvivalProteins[randomization.vector[]==1,]
sample.to.predict<-SurvivalProteins[randomization.vector[]==0,]
mtry<-3
ntree<-1000
node.size<-3

tree.that.predicts<-rsf(Survival,STATUS~STAGE+GENDER+
+HISTOLOGY+RESECTION+OTHER.DISEASE+TREATMENT+SMOKING+
+EGFR.FISH+EGFR.MUTATION+KRAS.MUTATION+
+EGFR.M+EGFR.N+EGFR.C+
+PTEN.M+PTEN.N+PTEN.C+
```
SUPPLEMENTARY INFORMATION CHAPTER 8

SUPPLEMENTARY FIGURE S1. ASSOCIATION BETWEEN GRANULAR PERK STAINING AND KRAS MUTATIONS

Panels A and B show representative examples of NSCLC tissue samples showing diffuse pERK staining in the nucleus and cytoplasm (A) or granular pERK staining in the cytoplasm (B). We observed an association between KRAS mutations and the presence of a typical granular pERK staining. In the graph (C) it can be observed that the proportion of patients having the granular staining pattern is higher in patients harboring a KRAS mutation than in patients having wild type KRAS status. This association was detected in a multiple testing procedure comprising tests on all other proteins, and may be called significant at a false discovery rate 30% or less (P=0.009). The molecular basis for this correlation between KRAS mutation and pERK granular staining pattern should be investigated in future fundamental research.
FIGURE S1

A  
DIFFUSE pERK STAINING

B  
GRANULAR pERK STAINING

C  
Granular staining pattern of pErk and KRAS mutation status

<table>
<thead>
<tr>
<th></th>
<th>NO GRANULAR pERK</th>
<th>GRANULAR pERK</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRAS WT</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>KRAS MUTANT</td>
<td>0%</td>
<td>100%</td>
</tr>
</tbody>
</table>

n=39  
n=13
APPENDIX B

SUPPLEMENTARY INFORMATION

CHAPTER 9

INTEGRATION OF GENE DOSAGE AND GENE EXPRESSION IN NON-SMALL CELL LUNG CANCER; IDENTIFICATION OF HSP90 AS POTENTIAL TARGET
## Table S1: Genes identified by ACE-it (359) ordered according to their relation with survival (Univariate cox regression analysis, raw p-values)

<table>
<thead>
<tr>
<th>ID</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Start</th>
<th>End</th>
<th>Clone</th>
<th>Gene name</th>
<th>Raw p-value</th>
<th>Adj p-value</th>
<th>Raw p-value</th>
<th>Adj p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>8984</td>
<td>BC032907</td>
<td>5</td>
<td>10405817</td>
<td>10406728</td>
<td>BC032907</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.003407233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9076</td>
<td>NM_016180</td>
<td>5</td>
<td>33906476</td>
<td>34020607</td>
<td>NM_016180</td>
<td>SLCA5A</td>
<td>0.000541</td>
<td>0.014974</td>
<td>0.005674666</td>
<td></td>
</tr>
<tr>
<td>9118</td>
<td>BC037711</td>
<td>5</td>
<td>36912618</td>
<td>37200061</td>
<td>BC037711</td>
<td>NPBBL</td>
<td>0.001968</td>
<td>0.030332</td>
<td>0.010061377</td>
<td></td>
</tr>
<tr>
<td>9178</td>
<td>U94910</td>
<td>5</td>
<td>43172632</td>
<td>43212250</td>
<td>NM_018038</td>
<td>SLC45A2</td>
<td>0.004104</td>
<td>0.045148</td>
<td>0.017059899</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>9156</td>
<td>I_935846</td>
<td>5</td>
<td>16190932</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9125</td>
<td>NM_016034</td>
<td>5</td>
<td>37415071</td>
<td>37788253</td>
<td>AK022353</td>
<td>RPL37</td>
<td>0.001968</td>
<td>0.030332</td>
<td>0.010061377</td>
<td></td>
</tr>
<tr>
<td>7121</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>9117</td>
<td>NM_015384</td>
<td>5</td>
<td>36912618</td>
<td>37100061</td>
<td>NM_015384</td>
<td>NIPBL</td>
<td>0.004916</td>
<td>0.049385</td>
<td>0.020917942</td>
<td></td>
</tr>
<tr>
<td>7122</td>
<td>I_928694</td>
<td>3</td>
<td>16169927</td>
<td>16170112</td>
<td>AK055675</td>
<td>RPL37</td>
<td>0.001207</td>
<td>0.021651</td>
<td>0.024471588</td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Chromosome</td>
<td>Start Position</td>
<td>End Position</td>
<td>Log2 Fold Change</td>
<td>p-Value</td>
<td>FDR</td>
<td>_q-Value</td>
<td>p-Value</td>
<td>FDR</td>
<td>q-Value</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
</tr>
<tr>
<td>7490</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7491</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7492</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7493</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7494</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7495</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7496</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7497</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7498</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7499</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7500</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7501</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7502</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7503</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7504</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7505</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7506</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7507</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7508</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7509</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7510</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7511</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7512</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7513</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7514</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7515</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7516</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7517</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7518</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7519</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7520</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7521</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7522</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7523</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7524</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7525</td>
<td>NM_138487</td>
<td>3</td>
<td>198157894</td>
<td>198159194</td>
<td>0.00274</td>
<td>0.030666</td>
<td>0.347508344</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SUPPLEMENTARY INFORMATION CHAPTER 9**
SUPPLEMENTARY INFORMATION CHAPTER 9
APPENDIX C

SUPPLEMENTARY INFORMATION

COLORED FIGURES
CHAPTER 3 - FIGURE 2

**Figures:**

- **A** Histology: microscopic image of a hematoxylin and eosin stained section showing squamous-cell carcinoma histology.
- **B** Immunohistochemistry of positive HER1/EGFR staining categorized as 3+ (EGFR clone 113, catalogue #MONX10177, 1:20 dilution, Monosan, Uden, The Netherlands).
- **C** Immunohistochemistry of positive phospho-Akt staining, categorized as 3+ in the nucleus and 2+ in the cytoplasm (p-Akt Ser 473 monoclonal antibody, catalogue #9277, 1:150 dilution, Cell Signaling, Beverly, MA, USA).
CHAPTER 5 - FIGURE 1

**FIGURE 1.** Sequence chromatograms for sample #3877909 showing a wild type sequence when DNA from frozen tissue was used as template for PCR, whereas when DNA extracted from the fixed and paraffin embedded (FPE) sample was used as template for PCR, a mutation in exon 20 could be observed, Gly779Ser (arrow). Re-analyzing this fixed and paraffin embedded sample revealed a wild type sequence.

CHAPTER 5 - FIGURE 2

**FIGURE 2.** (A) Sequencing EGFR from gDNA and RNA in sample #8446712 showing the single nucleotide polymorphism Gln787Gln and (B-C) the EGFR point mutations Ser768Ile and Leu861Gln. Whereas the wildtype to mutant allele ratio in the polymorphism is comparable in the chromatograms form gDNA versus RNA (A - arrows), both EGFR point mutation peaks are elevated in the chromatograms from RNA compared to the chromatograms from gDNA (B and C - arrows).
COLORED FIGURE ON BACK SIDE OF THIS PAGE
COLORED FIGURE ON BACK SIDE OF THIS PAGE
CHAPTER 6 - FIGURE 1

**FIGURE 1.** (A) Frozen tumor section negative for EGFR gene amplification showing 2-3 copies of the EGFR gene by Chromogenic In Situ Hybridization and (B) by Fluorescent In Situ Hybridization. (C) Frozen tumor section showing high EGFR gene copy number, detected by Chromogenic In Situ Hybridization. (D) A similar result is found with Fluorescent In Situ Hybridization, showing 6-10 dots per nucleus defined as high polysomy. Surrounding lymphocytes show two copies. (E) Paraffin tumor section positive for EGFR gene amplification showing EGFR gene clusters by Chromogenic In Situ Hybridization and (F) by Fluorescent In Situ Hybridization.
CHAPTER 7 – FIGURE 2

**FIGURE 2. Activation of Akt and Erk pathways in cells expressing mutant YFP-EGFR-ICD Del746.**

(A) Panels show representative images (400X) of MCF-7 cells expressing YFP-EGFR-ICD wild type, Del746 or VM68R, analyzed by immunofluorescence to detect phosphorylated Akt (pAkt). Phosphorylation of endogenous Akt was only detected in cells expressing the Del746-bearing protein. (B) A similar analysis was carried out to detect phosphorylated ERK (pERK). Only cells expressing YFP-EGFR-ICD Del746 contained detectable levels of endogenous pERK. Exposure time is indicated inside the panels. DNA was counterstained with Hoechst. (C) Images (400X) illustrate two morphological characteristics of Akt phosphorylation in cells expressing YFP-EGFR-ICD Del746. On one hand, pAkt showed a preferential localization to membrane ruffles, and often accumulated at the tip of lamellipodial protrusions (arrowhead). On the other hand, cells expressing high (cell#1) or low (cell#2) levels of YFP-EGFR-ICD Del746, often contained similar levels of pAkt.

CHAPTER 7 – FIGURE 3

**FIGURE 3. Evaluating erlotinib sensitivity of EGFR mutants in the context of YFP-EGFR-ICD.**

(A) Representative examples of MCF-7 cells expressing YFP-EGFR-ICD wild type, Del746 or Del746/T790M (green), stained for endogenous phosphorylated Akt (red). Cells were treated for 20 hours with the indicated concentration of erlotinib. The different response of each EGFR variant to erlotinib treatment is readily visualized by immunofluorescence. YFP-EGFR-ICD wild type does not induce Akt phosphorylation, and relocates into thick cytoplasmic fibrils at 10 μM erlotinib. One thousand-fold lower concentration of the drug (10 nM) inhibited Del746-induced Akt phosphorylation, and caused fibrilar relocation of the ectopic protein. The double mutant Del746/T790M did not form fibrils and induced Akt phosphorylation even in the presence of 10 μM erlotinib. (B) Images show that YFP-EGFR-ICD Del746 (green) remains phosphorylated at Y1092 (red) after relocating into fibrils in the presence of erlotinib.
SUPPLEMENTARY INFORMATION COLORED FIGURES

COLORED FIGURE ON BACK SIDE OF THIS PAGE
CHAPTER 7 – FIGURE 4

FIGURE 4. Testing the kinase activity and erlotinib sensitivity of uncommon EGFR mutants using the YFP-EGFR-ICD assay. (A) Partial amino acid sequence of EGFR exon 20 and exon 21 illustrating the location of the mutations examined (red letters). (B) Kinase activity and erlotinib sensitivity of different exon 20 mutations. Graph shows that autophosphorylation levels are lower for T790M (white diamonds) than for S768I (white squares) or Ins770SV (black triangles). Low expression levels hampered the accurate evaluation of Ins774HV (black circles). Images show that YFP-EGFR T790M did not effectively induce phosphorylation of endogenous Akt in MCF-7 cells, and did not relocate into fibrils upon erlotinib treatment. S768I-induced pAkt was inhibited by 100 nM erlotinib and the ectopic protein relocated into fibrils at 1 μM. The phosphorylation of Akt induced by exon 20 insertions was only inhibited at 10 μM erlotinib. This drug concentration also induced relocation of YFP-EGFR-ICD Ins770SV into fibrils. (C) Kinase activity and erlotinib sensitivity of different exon 21 mutations. Graph shows that the common L858R mutation confers higher autophosphorylation levels to YFP-EGFR-ICD than P848L and A859T. Images show that, unlike L858R, these uncommon exon 21 mutants did not induce phosphorylation of endogenous Akt. Erlotinib blocked L858R-induced pAkt at 10nM, and caused relocation of the ectopic protein into fibrils at 100 nM. Both effects were readily abrogated by the TKI-resistant mutation T790M. In all cases, data corresponding to one experiment are shown. Each EGFR mutant was tested at least twice with similar results.
CHAPTER 8 – FIGURE 2

FIGURE 2. Representative immunohistochemical staining pattern for several of the markers analyzed. (A) Examples of positive stainings of total EGFR on the membrane, (B) pCMET.1003 on the membrane (C), CMET.1349 in the nucleus, (D) E-CADHERIN on the membrane, (E) tumor cells negative for PTEN with positive stromal staining and (F) positive pSTAT3 staining in both the cytoplasm and nucleus.
CHAPTER 9 – FIGURE 1

**FIGURE 1.** Percentage of called gains and losses and their effect on gene expression in 32 NSCLC patients. Summary plot for called gains and losses in 32 resected NSCLC patients with DNA copy number changes indicated in grey. Positive values indicate the percentage of samples found with a gain. Negative values indicate the percentage of samples harboring a loss at the specified chromosome location. Genes in specified regions affected by copy number gain are indicated in green and genes affected by copy number loss are indicated in red. A selection of affected genes is indicated. The full list of 359 affected transcripts can be found in supplementary table S1.

CHAPTER 9 – FIGURE 4

**FIGURE 4.** Sensitivity of a panel of wild type EGFR-expressing NSCLC cell lines to the Hsp90 inhibitor 17-AAG. (A-D) time- and dose-dependent inhibition of the in vitro growth of H460, H157, H441, and A549 NSCLC cell lines following exposure to 17-AAG. Cells were seeded at $10^5$/well, and viable cell number was determined on subsequent days as described in Methods. 17-AAG concentrations at (H441) or above (A549, H460 & H157) 30 nM uniformly resulted in time-dependent loss of cell viability. The IC$_{50}$ value of 17-AAG (continuous exposure for 72 h) for each cell line is as follows: H460 = 30 nM, H157 = 15 nM, H441 = 8 nM, and A549 = 20 nM.
NEDERLANDSE SAMENVATTING
NEDERLANDSE SAMENVATTING VOOR NIET-INGEWIJDEN

In gezonde personen is er een goede balans tussen cellen die delen en cellen die doodgaan. In sommige gevallen wordt deze balans verstoord en delen cellen sneller dan ze doodgaan. Door deze ongecontroleerde deling van cellen ontstaan ophoping van cellen op specifieke plaatsen in het lichaam, een tumor, ook wel kanker genoemd. Kanker is een van de grootste doodsoorzaken in de westerse wereld. In Nederland is het, na hart- en vaatziekten, de grootste doodsoorzaak.

De deling van cellen in gezonde personen is een nauwkeurig gereguleerd proces, deels vastgelegd in de kern van de cel. In de kern van de cel bevindt zich het erfelijk materiaal, DNA, dit ligt bij mensen verspreid over 22 chromosomen. Delen van het DNA (genen) worden afgeschreven als RNA. Het RNA vormt de code om eiwitten te produceren. Eiwitten zijn uiteindelijk verantwoordelijk voor het uitoefenen van verschillende processen in de cel. Wanneer er een fout zit in het DNA, het RNA of in een eiwit kan het hele proces van bijvoorbeeld celdeling verstoord raken, wat vervolgens kan resulteren in kanker.

Longkanker is het type kanker dat verantwoordelijk is voor het grootste aantal kankergerelateerde sterftegevallen in de westerse wereld en in Azië. Patiënten kunnen worden behandeld door het wegsnijden van de tumor, soms wordt de hele long verwijderd. Er zijn echter vele gevallen waarin de toestand van de patiënt het niet toelaat een operatie te ondergaan. Ofwel omdat de patiënt niet fit genoeg is om met één long door te leven, en dus zou kunnen overlijden aan de operatie. Ofwel, omdat de longtumor in zo’n mate is doorgegroeid naar andere delen van het lichaam (is gemetastaseerd) dat het verwijderen van de tumorcellen niet mogelijk is door een operatie. In die gevallen wordt de patiënt behandeld met radiotherapie of chemotherapie. Deze vormen van therapie zorgen ervoor dat de deling van cellen wordt verstoord en dus de groei van de tumor wordt gehinderd. Echter de deling van gezonde cellen, bijvoorbeeld in het haar of de huid, worden met deze vorm van therapie ook verstoord wat leidt tot nare bijwerkingen zoals misselijkheid en haaruitval. Bovendien zijn deze vormen van therapie niet altijd effectief of wordt de tumor resistent tegen de therapie.

Om de effectiviteit van therapie te verhogen en de bijwerkingen van chemo- en radiotherapie te verminderen is de laatste jaren veel aandacht besteed aan nieuwe therapeutische benaderingen. Een nieuwe benadering is ‘targeted therapy’. Daarbij wordt een geneesmiddel ontwikkeld, die kan aangrijpen op structuren specifiek aanwezig op tumorcellen en niet, of in mindere mate, aanwezig op gezonde cellen.

Een voorbeeld van een geneesmiddel dat aangrijpt op zo’n specifieke structuur is een tyrosine kinase remmer van de epidermale groei factor receptor (EGFR). EGFR is in overmaat aanwezig op
tumorcellen van verschillende tumoren, waaronder op die van longtumoren. EGFR is een eiwitstructuur aanwezig op de buitenkant van cellen, en is in staat signalen door te geven naar de binnenkant van de cel om daar verschillende processen te reguleren. In gezonde cellen is EGFR verantwoordelijk voor het doorgeven van signalen voor een goede huishouding in de cel en tevens is dit eiwit belangrijk bij de embryonale ontwikkeling.

Wanneer patiënten met longkanker worden behandeld met EGFR tyrosine kinase remmers (EGFR TKIs) respondeert slechts 10% van de patiënten met verminderde groei of verdwijning van de tumor. Deze patiënten vertonen overigens weinig bijwerkingen van de therapie. Aangezien deze therapie dus maar bij een klein deel van de patiënten aanslaat is het belangrijk te onderzoeken welke eigenschappen van de tumor of de patiënt zijn gerelateerd aan een goede respons, om op deze manier het geneesmiddel zo effectief mogelijk te kunnen toepassen. Verschillende studies hebben uitgewezen dat patiënten met een goede respons voornamelijk bestonden uit vrouwelijke niet-rokers van Aziatische afkomst met het adenocarcinoma subtype. Op moleculair niveau, werd in het jaar 2004 ontdekt dat specifieke mutaties in het EGFR gen van tumorcellen was gerelateerd aan een goede respons op de EGFR tyrosine kinase remmers (EGFR TKIs). EGFR mutaties zouden dus gebruikt kunnen worden, in combinatie met de klinische eigenschappen van de patiënt als ‘biomarkers’, om de respons van de EGFR TKIs te voorspellen.

Helaas zijn weinig situaties helemaal zwart-wit en bestaat er altijd een grijs gebied. Zo is de correlatie tussen respons op EGFR TKIs en de aanwezigheid van EGFR mutaties niet 100%. Patiënten zonder mutaties kunnen alsnog reageren en ook patiënten met mutaties hebben de kans niet te reageren. Om deze reden wordt veel onderzoek verricht naar het vinden van additionele ‘biomarkers’ om respons of resistentie voor EGFR TKIs te voorspellen. Zo hebben verschillende onderzoeksgroepen aangetoond dat het aantal EGFR kopieën van het DNA een betere biomarker is voor respons op EGFR TKIs. Verder is gebleken dat mutaties in een ander gen, KRAS, juist verantwoordelijk zijn voor primaire resistentie tegen EGFR TKIs. Bovendien vertonen sommige patiënten na een initiële respons op een EGFR TKI, secundaire resistentie tegen dit geneesmiddel. In 50% van de gevallen kan dit worden verklaard door een tweede mutatie in het EGFR gen, de zogenaamde T790M mutatie. Verder is recentelijk gebleken dat een verhoogd aantal kopieën van het eiwit MET hiervoor ook verantwoordelijk kan zijn.

Uit bovenstaande feiten blijkt dat er goede vooruitgang wordt geboekt met het vinden van biomarkers voor voorspelling van respons op EGFR TKIs. Echter, hiermee worden minder dan 10% van de patiënten geholpen. Het is belangrijk dat er ook biomarkers worden gevonden die kunnen voorspellen hoe een patiënt op bijvoorbeeld chemotherapie zal reageren. Ook is het van belang biomarkers te ontdekken die kunnen voorspellen wat de prognose van de patiënt is, om aan de hand daarvan therapie aan te kunnen passen.
Een krachtige techniek om nieuwe genen te identificeren die belangrijk kunnen zijn in verschillende ziektes is 'microarray'. Bij deze techniek wordt het DNA of RNA van tumorcellen gelabeld met bijvoorbeeld een rode kleursof en DNA of RNA van een referentie (bijvoorbeeld een pool van tumor cellijnen) met een groene kleurstof. Vervolgens wordt het mengsel van deze gelabelde DNA of RNA samples op een microarray gepipetteerd. De microarray gebruikt voor ons onderzoek, bestaat uit een microscoop glasje waarop chemische structuren zijn geprint die overeenkomen met de genen op onze chromosomen. Het DNA of RNA zal specifiek aan deze gen-eigen chemische structuren binden onder experimentele condities. Vervolgens kan worden gemeten hoe sterk de rode (tumor) of groene (referentie) kleurstof heeft gebonden en zo kan de hoeveelheid DNA of RNA van alle genen worden gemeten. Wanneer men bijvoorbeeld wilt onderzoeken welke genen verantwoordelijk zijn voor een goede respons op een specifieke therapie, of welke genen voorspellen zijn voor een lange overleving, kan men deze techniek toepassen. Zo kan de genexpressie (RNA) of het aantal DNA kopieën worden gemeten in twee groepen patiënten met verschillende respons op een bepaald geneesmiddel. Wanneer een groep patiënten die goed heeft geraagd op de therapie significante verschillen vertoond, in RNA expressie of aantal DNA kopieën, met patiënten die niet hebben gereageerd kunnen deze genen of delen DNA verantwoordelijk zijn en nader worden onderzocht.

Het onderzoek beschreven in dit proefschrift omvat een nader onderzoek naar biomarkers in longkanker voor voorspelling van respons en prognose. In hoofdstuk 1 wordt een algemene inleiding gegeven over longkanker en de verschillende behandelmethode. Er wordt hier tevens dieper ingegaan op EGFR en het gebruik van EGFR tyrosine kinase remmers in de kliniek, tevens wordt de microarray techniek geïntroduceerd en de verschillende studies reeds uitgevoerd met deze techniek worden samengevat.

In hoofdstuk 2 wordt een klinische studie beschreven waarbij het middel Erlotinib (een EGFR TKI) wordt getest op longkanker patiënten. Van deze patiënten is ook bepaald of EGFR of KRAS mutaties aanwezig waren. Deze studie heeft bevestigd dat er een significante relatie bestaat tussen respons op de EGFR TKI en de aanwezigheid van EGFR en KRAS mutaties. Zoals eerder genoemd bestaan er altijd uitzonderingen, een van deze uitzonderingen wordt in detail beschreven in hoofdstuk 3. Hier wordt een mannelijke, westerse patiënt met longkanker beschreven die is behandeld met een EGFR TKI. Deze patiënt heeft eerder gerookt en het squamous cel carcinoma subtype en geen EGFR mutaties. Hoewel dit precies alle karakteristieken zijn die niet gerelateerd zijn aan respons op EGFR TKIs, vertoonde deze patiënt een goede respons op Erlotinib. Dit hoofdstuk benadrukt dus dat er additionele biomarkers gevonden moeten worden om respons op EGFR TKIs te voorspellen.

Bij patiënten met longkanker kan het voorkomen dat er meerdere tumoren binnen een persoon ontstaan. De diagnose of het gaat om verschillende primaire tumoren of om een metastase van
een primaire tumor, kan belangrijk zijn bij de keuze voor de juiste behandelwijze. In hoofdstuk 4 worden drie patiënten beschreven waarin verschillende tumoren aanwezig waren. In dit hoofdstuk wordt beschreven hoe het gebruik van EGFR en KRAS mutatie analyse in combinatie met de microarray techniek kunnen bijdragen aan het bepalen van de moleculaire relatie tussen verschillende tumoren binnen één patiënt.

Er wordt momenteel veel aandacht besteed aan het vinden van de meest optimale techniek om EGFR en KRAS mutaties te bepalen. In hoofdstuk 5 wordt een studie beschreven waarin de meest gebruikte techniek, PCR en sequencen, wordt vergeleken in monsters die verschillend gefixeerd en opgeslagen zijn geweest. Ook wordt vergeleken hoe de analyse verschilt met het gebruik van DNA of RNA als uitgangsmateriaal. Er wordt geconcludeerd dat het ideale uitgangsmateriaal RNA is, dat afkomstig is uit een sample die in de stikstof heeft opgeslagen gelegen. Verder worden in dit hoofdstuk aanwijzingen gegeven voor de mutatie analyse van EGFR en KRAS wanneer men niet beschikt over dit type samples.

Zoals eerder genoemd bleek naast EGFR mutaties, ook het aantal EGFR kopieën in het DNA van tumorcellen een goede voorspeller voor respons op EGFR TKIs te zijn. In hoofdstuk 6 beschrijven wij een vergelijkingsstudie naar twee verschillende technieken om het aantal EGFR kopieën in patiënten aan te tonen. We vergeleken de veel gebruikte techniek, fluorescente in situ hybridisatie (FISH) met een chromogene in situ hybridisatie (CISH). Wij tonen hier aan dat CISH een goed alternatief kan zijn voor FISH om het aantal EGFR kopieën te bepalen, CISH heeft bovendien een aantal praktische voordelen ten opzichte van FISH.

Er zijn verschillende EGFR mutaties gevonden in patiënten maar niet alle typen mutaties laat een even goede relatie zien met respons op EGFR TKIs. Zo is er bijvoorbeeld de eerder genoemde T790M mutatie die juist samengaat met secundaire resistentie. In hoofdstuk 7 beschrijven wij een techniek die wij hebben opgezet om op een snellere manier verschillende typen mutaties te onderzoeken op hun gevoeligheid voor EGFR TKIs (of andere geneesmiddelen).

Het feit dat EGFR een belangrijke rol speelt bij de voorspelling op EGFR TKIs bij longkanker patiënten benadrukt dat dit een belangrijk eiwit is in dit type tumoren. EGFR werkt samen met een groot aantal andere eiwitten om signalen door te geven aan de kern van de cel. In hoofdstuk 8 beschrijven wij een studie waarin we in monsters van longkanker patiënten hebben gekeken naar verschillende EGFR gerelateerde eiwitten. We hebben onder andere EGFR mutaties, KRAS mutaties en het aantal EGFR kopieën in het DNA bepaald. Wij beschrijven een statistisch model om de prognose van de patiënten te voorspellen aan de hand van de verschillende EGFR gerelateerde eiwitten. Wij vergelijken het systeem dat momenteel wordt gebruikt om prognose van de patiënt te bepalen, het zogenaamde pTNM systeem, met het gebruik van de EGFR gerelateerde moleculaire markers. Wij tonen aan dat met de moleculaire markers, een 7-10% betere nauwkeurigheid wordt bereikt om prognose van longkanker patiënten te bepalen, dan
wanneer het pTNM systeem wordt gebruikt. Hoewel dit geen schokkende verbetering is, tonen we aan dat een combinatie analyse als hier beschreven een geschikte methode kan zijn om prognose van patiënten beter te kunnen voorspellen.

Zoals blijkt uit vele studies, is de rol van EGFR in longkanker duidelijk aangetoond. Desondanks, blijft er een groot aantal patiënten waarbij niet bekend is wat de drijvende kracht achter de vorming van de tumor is geweest. We hebben de eerder beschreven microarray techniek gebruikt om additionele genen te identificeren die een rol zouden kunnen spelen in longkanker. Deze studie staat beschreven in hoofdstuk 9. We laten zien dat er 359 genen zijn die mogelijk een belangrijke rol te spelen in longkanker en we focussen in dit hoofdstuk op een van deze genen, HSP90. We tonen aan dat het HSP90 gen een duidelijke correlatie heeft met overleving van patiënten. Patiënten bij wie dit gen in mindere mate aanwezig was, overleefde significant langer dan patiënten waarbij dit gen in hogere mate aanwezig was. Dit kan een aanwijzing zijn dat remming van dit gen het leven van patiënten met longkanker zou kunnen verlengen.

Overigens is er al een 'targeted' geneesmiddel tegen dit gen, deze studie benadrukt dus dat dit geneesmiddel ook actief zou kunnen zijn in de behandeling van longkanker patiënten.

Tot slot geven we in hoofdstuk 10 een samenvatting van de studies beschreven in dit proefschrift. Vervolgens bediscussiëren we verschillende aspecten in de context van bestaande literatuur en geven daarmee aanwijzingen voor nader onderzoek in de toekomst.

Het onderzoek omschreven in dit proefschrift laat zien dat er de laatste jaren grote vooruitgang is geboekt met de identificatie van biomarkers bij longkanker patiënten. Vooral in de EGFR targeted therapy hebben EGFR en KRAS mutaties gebleken goede markers te zijn gerelateerd aan respons op EGFR TKIs, hoewel uitzonderingen altijd voor blijven komen. Wij laten in dit proefschrift verder zien dat de microarray techniek een goede techniek is om tot een selectie genen te komen die vervolgens nader bestudeerd zullen moeten worden om hun rol in longkanker uitvoeriger te bestuderen. Met de opkomst van verschillende technieken voor het screenen van zowel DNA, mRNA, niet coderend RNA (microRNAs) als ook eiwitten in tumorcellen, hopen wij in de toekomst een betere kijk te hebben op de moleculaire mechanismen verantwoordelijk voor het ontstaan, aanhouden en uitzaaien van tumoren. Door een juist gebruik van deze verschillende technieken al dan niet in combinatie met elkaar, en door samenwerkingen in de hele wereld, zullen we hopelijk in de toekomst het leven van patiënten met kanker een stuk dragelijker, en in het ideale geval, kankervrij kunnen maken.
CURRICULUM VITAE
CURRICULUM VITAE


CURRICULUM VITAE

Mariëlle Gallegos Ruiz was born on 18 oktober 1981 in Utrecht. In 1998 she obtained her HAVO diploma at College de Klop in Utrecht. Subsequently, she started her study Medical Microbiology at the Hogeschool van Utrecht. During this study she completed two internships. The first internship was performed at the diagnostic laboratory of medical microbiology from the st. Antonius Hospital, in Nieuwegein. She worked in the routine diagnostics under supervision of ing M.M.G de Goeij, ing D. van Ballegoy en dr G.P. Voorn. The second internship was fulfilled at Kiwa Water research in Nieuwegein. During this internship she developed a quick and sensitive detection method to detect *Legionella* in water, under supervision of ir B.A. Wullings en prof dr ir D. van der Kooij. In 2002 she graduated, after which she started a Master in Bio-Medical Sciences at the Vrije Universiteit in Amsterdam. During this study, again two internships were fulfilled. The first internship was performed at the VU University Medical Center, department of Medical Oncology under supervision of drs D. Fontijn en prof dr E. Boven. During this internship she investigated the role of basic fibroblast growth factor in melanoma cells. Her second internship was performed at the Paterson Institute for Cancer Research in Manchester, United Kingdom. Under supervision of dr L. Fairbairn en dr D. Gilham, she performed a functional analysis of Carcino Embryonic Antigen, transduced monocytes. Her literature study was written about mechanisms of action and resistance of monoclonal antibody therapy of B cell Non-Hodgkin’s Lymphoma. This was performed under supervision of dr M.J. Kersten en prof dr M.H.J. van Oers from the department of hematology from the Academic Medical Center in Amsterdam. This study is published in the “Nederlands Tijdschrift voor Hematologie” (nr5; 2005). She graduated on her Master Biomedical Sciences in 2004. Next, in October of 2004, she started as a research technician at the department of medical oncology of the VU University Medical Center in Amsterdam. After one year, she changed this position to work on the PhD project described in this thesis, entitled: “Therapeutic targets and biomarkers in lung cancer”. At present, she is completing additional microarray analyses related to the research described in this thesis. Furthermore, she is applying for financial support in order to get the opportunity to investigate various genes, described in chapter 9 of this thesis, and their role in lung cancer.
LIST OF PUBLICATIONS


DANKWOORD
DANKWOORD

Mijn naam mag dan misschien op de voorkant van dit boekje staan, maar dat betekent natuurlijk niet dat ik als enige verantwoordelijk ben voor al het werk dat is geleverd om dit boekje in elkaar te draaien. Daarom wil ik in dit hoofdstuk aandacht besteden aan alle personen die hier, direct of indirect, bij betrokken zijn geweest.

Allereerst wil ik mijn promotor prof dr Giuseppe Giaccone hartelijk bedanken. Dear Beppe, thank you so much for giving me the opportunity to work in your department and to give me the trust to start this PhD project. You have always encouraged me to get most out of myself, which I very much appreciate. Although I had to spend the last year without your physical presence, you were still always there for me when I needed your opinion. This is something I am very grateful for! I am also very thankful for all your help regarding my future.

Daarnaast wil ik hartelijke dank betuigen aan mijn copromotor dr J.A. Rodriguez. Dear Josean, I am highly thankful to you for everything that you have thought me. Every moment of the day you were there for me, you always came with bright ideas which has very much inspired me. I thought it was a big miss you left our department, although I am of course happy for you with your nice new position in Spain. Though a few thousand kilometers away, you were always there for me when I needed you. Thank you very much!

Aangezien mijn twee begeleiders fysiek niet meer aanwezig waren het laatste jaar, was er iemand nodig om een oogje in het zeil te houden. Prof dr G.J. Peters, Frits, heel erg bedankt dat je deze taak hebt overgenomen. Ik ben je erg dankbaar voor alle tijd die je altijd voor me had. Elk moment van de dag kon ik wel binnenlopen, met welke vraag dan ook. Dit heb ik altijd enorm gewaardeerd!

Er werd voor mij een tweede promotor aangesteld, prof dr E.F. Smit. Beste Egbert, ook jou wil ik enorm bedanken voor de goede begeleiding. Ik heb heel veel gehad aan je altijd snelle en kritische feedback! Ook wil ik je hartelijk bedanken voor de hulp bij mijn toekomstplannen.

Ook werd een tweede copromotor aangesteld, dr B. Ylstra. Beste Bauke, heel erg veel dank voor alle tijd die je hebt gestoken in het onderzoek. Ook bij jou kon ik altijd binnenlopen, en je altijd snelle en kritieke commentaar heb ik enorm gewaardeerd!
Alle leden van de leescommissie; prof dr G.A. Meijer, prof dr G.J. Peters, dr A.C. Dingemans, prof dr HJM Groen, dr L van ’t Veer en prof dr E.E. Voest, wil ik ook hartelijk danken voor hun tijd en aandacht voor dit proefschrift.

Een heel belangrijk persoon in dit onderzoek is Karijn Floor. Beste Karijn, zonder jou had ik pas de helft van de resultaten gehad! Ik vond het enorm fijn om met je samen te werken, altijd had je alles op orde, je was netjes, snel en nauwkeurig. Je bent echt een geweldige analist! Ik vond het dan ook jammer dat je tijd erop zat, maar ik ben erg blij dat je een leuke nieuwe baan hebt gevonden. Leuk dat je straks aan mijn zijde staat op het podium!

Simone heel erg bedankt voor het inwerken in het begin, en voor het mij altijd maar weer moeten uitleggen hoe je het makkelijkst een moeilijke berekening oploost;-). Ook bedankt voor de gezelligheid als kamergenootje, jammer dat je verhuist, een deur verder maakt al een hoop verschil! Gelukkig heb ik er leuke kamergenootjes voor teruggekregen! Kaamar en Leonie, al was het maar kort, bedankt voor de gezelligheid en heel veel succes met jullie onderzoek!

Alle gasten en studenten waarmee ik tijdens mijn onderzoek heb gewerkt wil ik ook bedanken! Laurens, ik vond het heel gezellig dat je er was, ik wens jou heel veel succes met je carrière! Fabrice, it was nice to have worked with you! Hope to visit you once in Marseille! Thijs, heel erg bedankt voor al jouw goede werk, wat staat beschreven in hoofdstuk 7. Ik wens je veel succes met je onderzoek in Leiden! Paolo, it was great to see how secure a medical doctor can work in the lab! You achieved very nice results! I wish you and your family all the best! Claudia, thanks for all your help with organizing the huge amount of TMA data! I wish you all the best for your future. Vanessa, you were a great student! You had your own project and only needed very little supervision. Wish you all the best in your career and also outside your career of course. Although Christmas is already over, the nice red nose Rudolph card still looks at me every day;-), I love that card! Francesca, many thanks for being such a nice roommate for a while. I am still loving the body lotion you gave me! Whenever I am in Italy I will call you to see where I can buy a new one, and to visit you of course! All the other Italian students; Silvia, Davide, Michele and Alberto, thanks for the nice time when you were in Amsterdam! Veerander, thanks for being such a nice roommate and I wish you lots of luck in your career! I would love to be guided through India by you!

De oude garde van de afdeling Oncologie; Maarten (bedankt voor de samenwerking in het begin), Marcel en Dennis (bedankt voor jullie gezelligheid en begeleiding tijdens mijn stage en
DANKWOORD

heel veel succes met jullie carrières). Agnieszka, (thank you so much for the nice chats we always had), Anne-marie (bedankt voor de tips rond voorbereidingen voor promoveren), Adrie, Anne, Astrid, Barbara, Bas, Eline, Eveline, Gerda, Jens (jij heel veel succes met je promotie op dezelfde dag!!! we gaan samen een leuk feestje geven, en heel erg bedankt voor het fijne samenwerken met voorbereidingen voor de promotie), Jennifer, Kasper, Michiel, Miriam, Olaf, Paul, Yuana. En de huidige garde; Irene, Auke, Clara, Elisa, thanks for the nice time we had during the AACR in LA and the nice time we will have in San Diego! En Laura, we gaan het vast ook gezellig hebben in San Diego! Ik wens jullie allemaal heel veel succes met het afronden van jullie onderzoek. Elisa has already finished her PhD (of course), I wish you lots of luck with all your research, I will not be surprised if I will see your name in the future as an intelligent professor! En de rest van de huidige garde; Henk D (bedankt voor al je computer en reference manager hulp!), Hester (bedankt voor de samenwerking voor hoofdstuk 4 en veel succes met jouw afrondingen!), Jelle (ook jij heel veel succes met je onderzoek, ik kijk uit naar de feestjes en etentjes in je grachtenpand:-)), Ida, Ietje (succes met CISH en FISH!), Jasmina (ook jou zie ik nog wel eens een goede professor worden;-)), Kees (volgens mij blijf jij altijd bij de "huidige garde" horen;-)), Maaike, Nienke, Richard H, Richard de H, Epie, Frank, Henk, Jan en Klaas wil ik bedanken. Verder Monique en Maartje, heel veel dank voor de hulp met geregeld rond mijn promotie! Maarten bedankt voor de financiële zaken. Sandra, ook jij bedankt voor alle gezellige praatjes, ik wens je je prille gezinnetje!

Vele andere afdelingen wil ik verder nog bedanken;
Mijn "adoptie" afdeling, de TP groep ben ik heel erg dankbaar! Beste Gerrit, heel erg bedankt dat ik altijd bij jullie meetings aanwezig mocht zijn, ik heb hier enorm veel van geleerd. Ook de rest van de TP groep wil ik hiervoor bedanken; Anke, Anne, Beatriz, Begoña, Danielle, Jeroen, Josien, Linda, Marianne, Mark, Meike, Mireille, Pien, Rebecca, Remond, Rinus, Sandra en Tineke, bedankt!
Ook natuurlijk de microarray faciliteit; François, ik ben je enorm dankbaar voor al het mooie werk dat je voor me hebt gedaan! Helaas hebben we dit werk niet meer in het boekje kunnen plaatsten, maar ik weet zeker dat we er een mooi artikel over kunnen schrijven, heel erg bedankt! Paul, veel dank voor het inwerken in het begin! Alan veel dank voor je hulp met GEO en heel veel succes met je nieuwe baan! Thijs, you still owe me something voor die keer dat ik voor niets uit mn bed ben gekomen;-)! Heel veel plezier op je wereldreis! Serge, bedankt voor de gezelligheid op het oude IMBL, veel succes met het afronden van jouw onderzoek, en veel plezier bij de microarray afdeling!
Dr Soria from Institute Gustave Roussy in Paris. Thank you very much for the collaboration on chapter 3.
DANKWOORD

De afdeling Moleculaire pathologie; Peter, Danielle, Duco, bedankt voor het succesvol overnemen van de EGFR mutatie analyse!

De afdeling Hematologie; Richard en Isabelle heel erg bedankt voor de samenwerking tijdens het project van Vanessa. En Isabelle, bedankt voor de tips rond mijn promotie!

De afdeling Reumatologie; Gerrit en Ruud (bedankt voor de gezellige praatjes en heel veel succes met je carrière!).

De afdeling Radiologie; Baukelien, heel erg bedankt voor de fijne samenwerking tijdens het project van Veerander!

Van de afdeling Fysiologie wil ik Cora hartelijk danken voor de hulp met de pholoidine kleuring.

De afdeling Longziekten; Remko heel erg bedankt voor je altijd snelle respons wanneer ik informatie nodig had uit de long database! Heel veel succes met het afronden van jouw onderzoek! Ook Erik, enorm bedankt voor de data die je mij hebt aangeleverd uit de database. Marieke, heel veel succes met je onderzoek!

Iedereen van het oude IMBL wil ik ook bedanken voor de gezelligheid in het lab toentertijd; Arjen, Hester, Fedor, Pauline, Hans, Janny, en Marjolein, bedankt!

De afdeling Pathologie; Helma enorm bedankt voor de TMA die je voor ons in korte tijd in elkaar hebt gezet en de altijd mooie coupes die je voor ons wist te snijden! Katrien en Erik ook jullie ben ik enorm dankbaar dat de deur altijd openstond. Een belletje was genoeg, en altijd stonden jullie voor me klaar om mijn coupes te beoordelen. Dit heb ik erg gewaardeerd! Erik ook nogmaals dank voor de uitnodiging te mogen spreken op de moleculaire pathologie bijeenkomst!

Verder wil ik ook heel graag Paul (agenda), Mark, Sjoerd (veel plezier op je reis!), Wessel en José heel hartelijk danken voor alle biostatistische discussies. Ik heb heel erg veel geleerd van jullie op dit vakgebied! Ook bedankt voor alle analyses die jullie voor mij hebben uitgevoerd, die echt aan mijn verstand te boven gingen!

Ook veel mensen buiten het lab wil ik heel erg bedanken;

Dirk heel erg bedankt voor het werk dat je hebt gedaan om mijn cover te maken tot wat hij is geworden, veel dank!

Al mijn lieve vriendjes en vriendinnen wil ik vooral bedanken voor de nodige sociale ontspanning. Bedankt voor de nachtelijke stapavonden, de bioscooppjes, etentjes, bedankt ook voor de salsa avondjes, de wijntjes op de bank, de biertjes in de kroeg, de girlnights (...), de vakanties naar de zon... Al deze ontspanning is toch een van de belangrijkste dingen in het leven, vooral in het leven van een promovendus:-)). Eef jij ook in het speciaal bedankt dat je van zover komt om me bij te staan op het podium! Love u!
Lieve Erik, heel erg bedankt voor de persoon die je bent en voor de fijne jaren samen. Vezels zijn... ? Hihi.. gewoon ff stom geintje tussendoor. Dikke kus en knuffel voor jou voor alles wat je voor me hebt betekend!

Lieve Robbie, Vivian, Timothy en Mitchel; bedankt voor de fijne avondjes altijd bij jullie thuis, lekker eten, relaxte films, leuke gesprekken. Bedankt!

Lieve Edisona, mi ke djis menshona bo nomber paso mi ta kontentu ku bo ta mi “schoonmoeder”. Bo ta un dushi persona! Duna John i Aida tambe un brasa pa mi!

Lieve familie, ooms, tantes, neefjes, nichtjes, Aart en Anna jullie ook bedankt voor de gezelligheid en alle interesse die jullie altijd hadden voor mijn werk!

Lief broertje, lieve zusjes en lieve Lloydiane, ook jullie heel erg bedankt voor de fijne personen die jullie zijn! Altijd fijn om jullie om me heen te hebben!

Lieve pap en mam, ik wil jullie heel erg bedanken voor alle mogelijkheden die jullie me altijd hebben gegeven om te kunnen doen wat ik wil, en te bereiken wat ik tot nu toe bereikt heb. Heel erg bedankt voor het fijne thuis wat ik altijd heb gehad, en nog steeds voel, bij jullie. Bedankt dat jullie er altijd, door dik en dun, voor me zijn. Ik hou van jullie!

Komo delaster hende mi ke yama danki na mi Engel(bertje). Dushi, danki pa ta e hende ku bo ta i ku semper bo ta kla pa yuda mi. Danki ku bo ta tene mi ku dos pia riba suela, sino masha dia lo mi a bula bai ☺. Mi ta masha kontentu ku e sintimentu ku nos tin huntu i mi ta apresiá e dushi kas ku bo ta dunami. Dushi mi stima bo hopi i mi ta sperma un futuro bunita ku bo.

Iedereen bedankt!!