(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 February 2009 (19.02.2009)

PCT

(10) International Publication Number
WO 2009/021673 A1

(51) International Patent Classification:
*C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/EP2008/006512

(22) International Filing Date: 7 August 2008 (07.08.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07114336.6 14 August 2007 (14.08.2007) EP

(71) Applicant *(for all designated States except US)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DELMAR, Paul** [FR/CH]; Hammerstrasse 89, CH-4057 Basel (CH). **KLUGHAMMER, Barbara** [DE/DE]; Im Adlergarten 11, 79618 Rheinfelden (DE). **LUTZ, Verena** [DE/DE]; Lindenschmitstrasse 25, 81371 Muenchen (DE). **MCLOUGHLIN, Patricia** [IE/CH]; Claragraben 115, CH-4057 BASEL (CH).

(74) Agent: **KUENG, Peter**; Grenzacherstrasse 124, CH-4070 Basel (CH).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**
— *of inventorship (Rule 4.17(iv))*

**Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(54) **Title:** PREDICTIVE MARKERS FOR EGFR INHIBITORS TREATMENT

(57) **Abstract:** The present invention provides biomarkers that are predictive for the response to treatment with an EGFR inhibitor in cancer patients. The markers are the genes GBAS, APOH, SCYL3, PMS2CL, PRODH, SERFlA, URG4A and LRRC31.

WO 2009/021673 A1

# Predictive markers for EGFR inhibitor treatment

The present invention provides biomarkers that are predictive for the response to treatment with an EGFR inhibitor in cancer patients

A number of human malignancies are associated with aberrant or over-expression of the epidermal growth factor receptor (EGFR). EGF, transforming growth factor α (TGF-α), and a number of other ligands bind to the EGFR, stimulating autophosphorylation of the intracellular tyrosine kinase domain of the receptor. A variety of intracellular pathways are subsequently activated, and these downstream events result in tumour cell proliferation in vitro. It has been postulated that stimulation of tumour cells via the EGFR may be important for both tumour growth and tumour survival in vivo.

Early clinical data with Tarceva™ (erlotinib), an inhibitor of the EGFR tyrosine kinase, indicate that the compound is safe and generally well tolerated at doses that provide the targeted effective concentration (as determined by preclinical data). Clinical phase I and II trials in patients with advanced disease have demonstrated that Tarceva™ has promising clinical activity in a range of epithelial tumours. Indeed, Tarceva™ has been shown to be capable of inducing durable partial remissions in previously treated patients with head and neck cancer, and NSCLC (Non small cell lung cancer) of a similar order to established second line chemotherapy, but with the added benefit of a better safety profile than chemo therapy and improved convenience (tablet instead of intravenous [i.v.] administration). A recently completed, randomised, double-blind, placebo-controlled trial (BR.21) has shown that single agent Tarceva™ significantly prolongs and improves the survival of NSCLC patients for whom standard therapy for advanced disease has failed.

Erlotinib (Tarceva™) is a small chemical molecule; it is an orally active, potent, selective inhibitor of the EGFR tyrosine kinase (EGFR-TKI).

Lung cancer is the major cause of cancer-related death in North America and Europe. In the United States, the number of deaths secondary to lung cancer exceeds the combined total deaths from the second (colon), third (breast), and fourth (prostate) leading causes of cancer deaths combined. About 75% to 80% of all lung cancers are NSCLC, with

approximately 40% of patients presenting with locally advanced and/or unresectable disease. This group typically includes those with bulky stage IIIA and IIIB disease, excluding malignant pleural effusions.

The crude incidence of lung cancer in the European Union is 52.5, the death rate 48.7 cases/100000/year. Among men the rates are 79.3 and 78.3, among women 21.6 and 20.5, respectively. NSCLC accounts for 80% of all lung cancer cases. About 90% of lung cancer mortality among men, and 80% among women, is attributable to smoking.

In the US, according to the American Cancer Society, during 2004, there were approximately 173,800 new cases of lung cancer (93,100 in men and 80,700 in women) and were accounting for about 13% of all new cancers.. Most patients die as a consequence of their disease within two years of diagnosis. For many NSCLC patients, successful treatment remains elusive. Advanced tumours often are not amenable to surgery and may also be resistant to tolerable doses of radiotherapy and chemotherapy. In randomized trials the currently most active combination chemotherapies achieved response rates of approximately 30% to 40% and a 1-year survival rate between 35% and 40%. This is really an advance over the 10% 1-year survival rate seen with supportive care alone.

Until recently therapeutic options for patients following relapse were limited to best supportive care or palliation. A recent trial comparing docetaxel (Taxotere®) with best supportive care showed that patients with NSCLC could benefit from second line chemotherapy after cisplatin-based first-line regimens had failed. Patients of all ages and with ECOG performance status of 0, 1, or 2 demonstrated improved survival with docetaxel, as did those who had been refractory to prior platinum-based treatment. Patients who did not benefit from therapy included those with weight loss of > 10%, high lactate dehydrogenase levels, multi-organ involvement, or liver involvement. Additionally, the benefit of docetaxel monotherapy did not extend beyond the second line setting. Patients receiving docetaxel as third-line treatment or beyond showed no prolongation of survival. Single-agent docetaxel became a standard second-line therapy for NSCLC. Recently another randomized phase III trial in second line therapy of NSCLC compared pemetrexed (Alimta®) with docetaxel. Treatment with pemetrexed resulted in a clinically equivalent efficacy but with significantly fewer side effects compared with docetaxel.

It has long been acknowledged that there is a need to develop methods of individualising cancer treatment. With the development of targeted cancer treatments, there is a particular interest in methodologies which could provide a molecular profile of the tumour

target, (i.e. those that are predictive for clinical benefit). Proof of principle for gene expression profiling in cancer has already been established with the molecular classification of tumour types which are not apparent on the basis of current morphological and immunohistochemical tests.

Therefore, it is an aim of the present invention to provide expression biomarkers that are predictive for response to EGFR inhibitor treatment in cancer patients.

In a first object the present invention provides an in vitro method of predicting the response of a cancer patient to treatment with an EGFR inhibitor comprising the steps: determining the expression level of at least one gene selected from the group consisting of GBAS, APOH, SCYL3, PMS2CL, PRODH, SERF1A, URG4A and LRR 31 in a tumour sample of a patient and comparing the expression level of the at least one gene to a value representative of an expression level of the at least one gene in tumours of a non responding patient population, wherein a higher expression level of the at least one gene in the tumour sample of the patient is indicative for a patient who will respond to the treatment.

The term "a value representative of an expression level of the at least one gene in tumours of a non responding patient population" refers to an estimate of the mean expression level of the marker gene in tumours of a population of non responding patients.

In a preferred embodiment, the expression level of the at least one gene is determined by microarray technology or other technologies that assess RNA expression levels like quantitative RT-PCR, or by any method looking at the expression level of the respective protein, e.g. immunohistochemistry (IHC). The construction and use of gene chips are well known in the art. see, U. S. Pat Nos. 5,202,231; 5,445,934; 5,525,464; 5,695,940; 5,744,305; 5,795, 716 and 1 5,800,992. See also, Johnston, M. Curr. Biol. 8:R171-174 (1998); Iyer VR et al., Science 283:83-87 (1999). Of course, the gene expression level can be determined by other methods that are known to a person skilled in the art such as e.g. northern blots, RT-PCR, real time quantitative PCR, primer extension, RNase protection, RNA expression profiling.

In a further preferred embodiment, the expression level of at least two genes is determined, preferably of at least three genes.

The genes of the present invention can be combined to biomarker sets. Biomarker sets can be built from any combination of biomarkers listed in Table 3 to make predictions about the effect of EGFR inhibitor treatment in cancer patients. The various biomarkers and

biomarkers sets described herein can be used, for example, to predict how patients with cancer will respond to therapeutic intervention with an EGFR inhibitor.

In a preferred embodiment, the marker gene in the tumour sample of the responding patient shows typically between 1.1 and 2.7 or more fold higher expression compared to a value representative of the expression level of the at least one gene in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene GBAS and shows typically between 1.4 and 2.7 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene GBAS in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene APOH and shows typically between 1.4 and 2.6 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene APOH in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene SCYL3 and shows typically between 1.3 and 1.8 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene SCYL3 in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene PMS2CL and shows typically between 1.2 and 1.5 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene PMS2CL in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene PRODH and shows typically between 1.5 and 3.0 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene PRODH in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene SERF1A and shows typically between 1.2 and 1.6 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene SERF1A in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene URG4 and shows typically between 1.1 and 1.3 or more fold higher expression in the tumour sample of the responding patient

compared to a value representative of the expression level of the gene URG4 in tumours of a non responding patient population.

In a preferred embodiment, the marker is gene LRRC31 and shows typically between 1.3 and 1.8 or more fold higher expression in the tumour sample of the responding patient compared to a value representative of the expression level of the gene LRRC31 in tumours of a non responding patient population.

The genes of the present invention can be combined to biomarker sets. Biomarker sets can be built from any combination of biomarkers listed in Table 3 to make predictions about the effect of EGFR inhibitor treatment in cancer patients. The various biomarkers and biomarkers sets described herein can be used, for example, to predict how patients with cancer will respond to therapeutic intervention with an EGFR inhibitor.

The term "gene" as used herein comprises variants of the gene. The term "variant" relates to nucleic acid sequences which are substantially similar to the nucleic acid sequences given by the GenBank accession number. The term "substantially similar" is well understood by a person skilled in the art. In particular, a gene variant may be an allele which shows nucleotide exchanges compared to the nucleic acid sequence of the most prevalent allele in the human population. Preferably, such a substantially similar nucleic acid sequence has a sequence similarity to the most prevalent allele of at least 80%, preferably at least 85%, more preferably at least 90%, most preferably at least 95%. The term "variants" is also meant to relate to splice variants.

The EGFR inhibitor can be selected from the group consisting of gefitinib, erlotinib, PKI-166, EKB-569, GW2016, CI-1033 and an anti-erbB antibody such as trastuzumab and cetuximab.

In another embodiment, the EGFR inhibitor is erlotinib.

In yet another embodiment, the cancer is NSCLC.

Techniques for the detection and quantitation of gene expression of the genes described by this invention include, but are not limited to northern blots, RT-PCR, real time quantitative PCR, primer extension, RNase protection, RNA expression profiling and related techniques. These techniques are well known to those of skill in the art see e.g. Sambrook J et al., Molecular Cloning: A Laboratory Manual, Third Edition (Cold Spring Harbor Press, Cold Spring Harbor, 2000).

Techniques for the detection of protein expression of the respective genes described by this invention include, but are not limited to immunohistochemistry (IHC).

In accordance with the invention, cells from a patient tissue sample, e.g. a tumour or cancer biopsy can be assayed to determine the expression pattern of one or more biomarkers. Success or failure of a cancer treatment can be determined based on the biomarker expression pattern of the cells from the test tissue (test cells), e.g., tumour or cancer biopsy, as being relatively similar or different from the expression pattern of a control set of the one or more biomarkers. In the context of this invention, it was found that the genes listed in table 3 are up-regulated i.e. show a higher expression level, in tumours of patients who respond to the EGFR inhibitor treatment compared to tumours of patients who do not respond to the EGFR inhibitor treatment. Thus, if the test cells show a biomarker expression profile which corresponds to that of a patient who responded to cancer treatment, it is highly likely or predicted that the individual's cancer or tumour will respond favourably to treatment with the EGFR inhibitor. By contrast, if the test cells show a biomarker expression pattern corresponding to that of a patient who did not respond to cancer treatment, it is highly likely or predicted that the individual's cancer or tumour will not respond to treatment with the EGFR inhibitor.

The biomarkers of the present invention i.e. the genes listed in table 3 are a first step towards an individualized therapy for patients with cancer, in particular patients with refractory NSCLC. This individualized therapy will allow treating physicians to select the most appropriate agent out of the existing drugs for cancer therapy, in particular NSCLC. The benefit of individualized therapy for each future patient are: response rates / number of benefiting patients will increase and the risk of adverse side effects due to ineffective treatment will be reduced.

In a further object the present invention provides a therapeutic method of treating a cancer patient identified by the in vitro method of the present invention. Said therapeutic method comprises administering an EGFR inhibitor to the patient who has been selected for treatment based on the predictive expression pattern of at least one of the genes listed in table 3. A preferred EGFR inhibitor is erlotinib and a preferred cancer to be treated is NSCLC.

**Short description of the figures**

Figure 1 shows the study design and

Figure 2 shows a scheme of sample processing.

**Experimental part**

Rationale for the Study and Study Design

Recently mutations within the EGFR gene in the tumour tissue of a subset of NSCLC patients and the association of these mutations with sensitivity to erlotinib and gefitinib were described (Pao W, et al. 2004; Lynch et al. 2004; Paez et al. 2004). For the patients combined from two studies, mutated EGFR was observed in 13 of 14 patients who responded to gefitinib and in none of the 11 gefitinib-treated patients who did not respond. The reported prevalence of these mutations was 8% (2 of 25) in unselected NSCLC patients. These mutations were found more frequently in adenocarcinomas (21%), in tumours from females (20%), and in tumours from Japanese patients (26%). These mutations result in increased in vitro activity of EGFR and increased sensitivity to gefitinib. The relationship of the mutations to prolonged stable disease or survival duration has not been prospectively evaluated.

Based on exploratory analyses from the BR.21 study, it appeared unlikely that the observed survival benefit is only due to the EGFR mutations, since a significant survival benefit is maintained even when patients with objective response are excluded from analyses. Other molecular mechanisms must also contribute to the effect.

Based on the assumption that there are changes in gene expression levels that are predictive of response / benefit to Tarceva™ treatment, microarray analysis was used to detect these changes

This required a clearly defined study population treated with Tarceva™ monotherapy after failure of 1st line therapy. Based on the experience from the BR.21 study, benefiting population was defined as either having objective response, or disease stabilization for $\geq 12$ weeks. Clinical and microarray datasets were analyzed according to a pre-defined statistical plan.

The application of this technique requires fresh frozen tissue (FFT). Therefore a mandatory biopsy had to be performed before start of treatment. The collected material was frozen in liquid nitrogen ($N_2$).

A second tumour sample was collected at the same time and stored in paraffin (formalin fixed paraffin embedded, FFPE). This sample was analysed for alterations in the EGFR signalling pathway.

The ability to perform tumour biopsies via bronchoscopy was a prerequisite for this study. Bronchoscopy is a standard procedure to confirm the diagnosis of lung cancer. Although generally safe, there is a remaining risk of complications, e.g. bleeding.

Rationale for Dosage Selection

Tarceva™ was given orally once per day at a dose of 150 mg until disease progression, intolerable toxicities or death. The selection of this dose was based on pharmacokinetic parameters, as well as the safety and tolerability profile of this dose observed in Phase I, II and III trials in heavily pre-treated patients with advanced cancer. Drug levels seen in the plasma of patients with cancer receiving the 150 mg/day dose were consistently above the average plasma concentration of 500 ng / ml targeted for clinical efficacy. BR.21 showed a survival benefit with this dose.

Objectives of the Study

The primary objective was the identification of differentially expressed genes that are predictive for benefit (CR, PR or SD ≥ 12 weeks) of Tarceva™ treatment. Identification of differentially expressed genes predictive for "response" (CR, PR) to Tarceva™ treatment was an important additional objective.

The secondary objectives were to assess alterations in the EGFR signalling pathway with respect to benefit from treatment.

**Study Design**

Overview of Study Design and Dosing Regimen

This was an open-label, predictive marker identification Phase II study. The study was conducted in approximately 26 sites in about 12 countries. 264 patients with advanced NSCLC following failure of at least one prior chemotherapy regimen were enrolled over a 12 month period. Continuous oral Tarceva™ was given at a dose of 150 mg/day. Dose reductions were permitted based on tolerability to drug therapy. Clinical and laboratory parameters were assessed to evaluate disease control and toxicity. Treatment continued until disease progression, unacceptable toxicity or death.

Tumour tissue and blood samples were obtained for molecular analyses to evaluate the effects of Tarceva™ and to identify subgroups of patients benefiting from therapy. The study design is depicted in figure 1.

Predictive Marker Assessments

Biopsies of the tumour were taken within 2 weeks before start of treatment. Two different samples were collected:

The first sample was always frozen immediately in liquid $N_2$.

The second sample was fixed in formalin and embedded in paraffin.

Snap frozen tissue had the highest priority in this study.

Figure 2 shows a scheme of the sample processing.

Microarray Analysis

The snap frozen samples were used for laser capture microdissection (LCM) of tumour cells to extract tumour RNA and RNA from tumour surrounding tissue. The RNA was analysed on Affymetrix microarray chips (HG-U133A) to establish the patients' tumour gene expression profile. Quality Control of Affymetrix chips was used to select those samples of adequate quality for statistical comparison.

Single Biomarker Analyses on Formalin Fixed Paraffin Embedded Tissue

The second tumour biopsy the FFPE sample was used to perform DNA mutation, IHC and ISH analyses as described below. Similar analyses were performed on tissue collected at initial diagnosis.

The DNA mutation status of the genes encoding EGFR and other molecules involved in the EGFR signalling pathway were analysed by DNA sequencing. Gene amplification of EGFR and related genes were be studied by FISH.

Protein expression analyses included immunohistochemical [IHC] analyses of EGFR and other proteins within the EGFR signalling pathway.

Response Assessments

The RECIST (Uni-dimensional Tumour Measurement) criteria were used to evaluate response. These criteria can be found under the following link:

http://www.eortc.be/recist/

Note that: To be assigned a status of CR or PR, changes in tumour measurements must be confirmed by repeated assessments at least 4 weeks apart at any time during the treatment period.

In the case of SD, follow-up measurements must have met the SD criteria at least once after study entry at a minimum interval of 6 weeks.

In the case of maintained SD, follow-up measurements must have met the SD criteria at least once after study entry with maintenance duration of *at least* 12 weeks.

Survival Assessment

A regular status check every 3 months was performed either by a patient's visit to the clinic or by telephone. All deaths were recorded. At the end of the study a definitive confirmation of survival was required for each patient.

**Methods**

RNA samples preparation and quality control of RNA samples

All biopsy sample processing was handled by a pathology reference laboratory; fresh frozen tissue samples were shipped from investigator sites to the Clinical Sample Operations facility in Roche Basel and from there to the pathology laboratory for further processing. Laser capture microdissection was used to select tumour cells from surrounding tissue. After LCM, RNA was purified from the enriched tumour material. The pathology laboratory then carried out a number of steps to make an estimate of the concentration and quality of the RNA.

RNases are RNA degrading enzymes and are found everywhere and so all procedures where RNA will be used must be strictly controlled to minimize RNA degradation. Most mRNA species themselves have rather short half-lives and so are considered quite unstable. Therefore it is important to perform RNA integrity checks and quantification before any assay.

RNA concentration and quality profile can be assessed using an instrument from Agilent (Agilent Technologies, Inc., Palo Alto, CA) called a 2100 Bioanalyzer®. The instrument software generates an RNA Integrity Number (RIN), a quantitation estimate (Schroeder, A., et al., The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol, 2006. 7: p. 3), and calculates ribosomal ratios of the total RNA sample. The RIN is determined from the entire electrophoretic trace of the RNA sample, and so includes the presence or absence of degradation products.

The RNA quality was analysed by a 2100 Bioanalyzer®. Only samples with at least one rRNA peak above the added poly-I noise and sufficient RNA were selected for further analysis on the Affymetrix platform. The purified RNA was forwarded to the Roche Centre

for Medical Genomics (RCMG; Basel, Switzerland) for analysis by microarray. 122 RNA samples were received from the pathology laboratory for further processing.

Target Labeling of tissue RNA samples

Target labeling was carried out according to the Two-Cycle Target Labeling Amplification Protocol from Affymetrix (Affymetrix, Santa Clara, California), as per the manufacturer's instructions.

The method is based on the standard Eberwine linear amplification procedure but uses two cycles of this procedure to generate sufficient labeled cRNA for hybridization to a microarray.

Total RNA input used in the labeling reaction was 10ng for those samples where more than 10ng RNA was available; if less than this amount was available or if there was no quantity data available (due to very low RNA concentration), half of the total sample was used in the reaction. Yields from the labeling reactions ranged from 20-180µg cRNA. A normalization step was introduced at the level of hybridization where 15µg cRNA was used for every sample.

Human Reference RNA (Stratagene, Carlsbad, CA, USA) was used as a control sample in the workflow with each batch of samples. 10ng of this RNA was used as input alongside the test samples to verify that the labeling and hybridization reagents were working as expected.

Microarray hybridizations

Affymetrix HG-U133A microarrays contain over 22,000 probe sets targeting approximately 18,400 transcripts and variants which represent about 14,500 well-characterized genes.

Hybridization for all samples was carried out according to Affymetrix instructions (Affymetrix Inc., Expression Analysis Technical Manual, 2004). Briefly, for each sample, 15µg of biotin-labeled cRNA were fragmented in the presence of divalent cations and heat and hybridized overnight to Affymetrix HG-U133A full genome oligonucleotide arrays. The following day arrays were stained with streptavidin-phycoerythrin (Molecular Probes; Eugene, OR) according to the manufacturer's instructions. Arrays were then scanned using a GeneChip Scanner 3000 (Affymetrix), and signal intensities were automatically calculated by GeneChip Operating Software (GCOS) Version 1.4 (Affymetrix).

Statistical Analysis

Analysis of the Affymetrix™ data consisted of four main steps.

Step 1 was quality control. The goal was to identify and exclude from analysis array data with a sub-standard quality profile.

Step 2 was pre-processing and normalization. The goal was to create a normalized and scaled "analysis data set", amenable to inter-chip comparison. It comprised background noise estimation and subtraction, probe summarization and scaling.

Step 3 was exploration and description. The goal was to identify potential bias and sources of variability. It consisted of applying multivariate and univariate descriptive analysis techniques to identify influential covariates.

Step 4 was modeling and testing. The goal was to identify a list of candidate markers based on statistical evaluation of the difference in mean expression level between "Responders" (patients with "Partial Response" or "Complete Response" as best response) and "Non Responders" (patients with "Stable Disease" or "Progressive Disease" as best response). It consisted of fitting an adequate statistical model to each probe-set and deriving a measure of statistical significance.

All analyses were performed using the R software package.

Step 1 : Quality Control

The assessment of data quality was based on checking several parameters. These included standard Affymetrix GeneChip™ quality parameters, in particular: Scaling Factor, Percentage of Present Call and Average Background. This step also included visual inspection of virtual chip images for detecting localized hybridization problems, and comparison of each chip to a virtual median chip for detecting any unusual departure from median behaviour. Inter-chip correlation analysis was also performed to detect outlier samples. In addition, ancillary measures of RNA quality obtained from analysis of RNA samples with the Agilent Bioanalyzer™ 2100 were taken into consideration.

Based on these parameters, data from 20 arrays were excluded from analysis. Thus data from a total of 102 arrays representing 102 patients was included in the analysis. The clinical description of these 102 patients set is reported in table 1.

Table 1: Description of clinical characteristics of patients included in the analysis.

| Variable | Value | n=102 |
|---|---|---|
| | | n (%) |
| Best Response | N/A | 16 (15.7%) |
| | PD | 49 (48.0%) |
| | SD | 31 (30.4%) |
| | PR | 6 (5.9%) |
| Clinical Benefit | NO | 81 (79.4%) |
| | YES | 21 (20.6%) |
| SEX | FEMALE | 25 (24.5%) |
| | MALE | 77 (74.5%) |
| ETHNICITY | CAUCASIAN | 65 (63.7%) |
| | ORIENTAL | 37 (36.3%) |
| Histology | ADENOCARCINOMA | 35 (34.3%) |
| | SQUAMOUS | 53 (52.0%) |
| | OTHERS | 14 (13.7%) |
| Ever-Smoking | NO | 20 (19.6%) |
| | YES | 82 (80.4%) |

Step 2 : Data pre-processing and normalization

The rma algorithm (Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data.* Nucl. Acids Res., 2003. **31**(4): p. e15) was used for pre-processing and normalization. The mas5 algorithm (AFFYMETRIX, *GeneChip® Expression: Data Analysis Fundamentals*. 2004, AFFYMETRIX) was used to make detection calls for the individual probe-sets. Probe-sets called "absent" or "marginal" in all samples were removed from further analysis; 5930 probe-sets were removed from analysis based on this criterion. The analysis data set therefore consisted of a matrix with 16353 (out of 22283) probe-sets measured in 102 patients.

Step 3 : Data description and exploration

Descriptive exploratory analysis was performed to identify potential bias and major sources of variability. A set of covariates with a potential impact on gene expression profiles was screened. It comprised both technical and clinical variables. Technical covariates included: Date of RNA processing (later referred to as batch), RIN (as a measure of RNA quality/integrity), Operator and Center of sample collection. Clinical covariates included: Histology type, smoking status, tumour grade, performance score, demographic data, responder status and clinical benefit status.

The analysis tools included univariate ANOVA and principal component analysis. For each of these covariates, univariate ANOVA was applied independently to each probe-set.

A significant effect of the batch variable was identified. In practice, the batch variable captured differences between dates of sample processing and Affymetrix chip lot. After checking that the batch variable was nearly independent from the variables of interest, the batch effect was corrected using the method described in Johnson et al., Biostat, 2007. 8(1): p. 118-127.

The normalized data set after batch effect correction served as the analysis data set in subsequent analyses.

Histology and RIN were two additional important variables highlighted by the descriptive analysis.

Step 4 : Data modeling and testing.

A linear model was fitted independently to each probe-set. Variables included in the model are reported in table 2. A linear model was fitted independently to each probe-set. Variables included in the model are reported in table 2. The model parameters were estimated by the maximum likelihood technique. The parameter corresponding to the "Response" variable (X1) was used to assess the difference in expression level between the group "responding" and "non responding" patients.

Table 2: Description of the variables included in the linear model.

| **Variable** | **Type** | **Values** |
|---|---|---|
| **gene expression** | **Dependent ($Y_{ip}$)** | **Normalized log2 intensity of probe-set i in patient p.** |
| Intercept | Overall mean ($\mu$) | |
| **Response** | **Predictor of interest (X1)** | **YES / NO** |
| Histology | Adjustment Covariate (X2) | ADENOCARCINOMA / SQUAMOUS / OTHERS |
| RACE | Adj. Cov. (X3) | ORIENTAL / CAUCASIAN |
| SEX | Adj Cov. (X4) | FEMALE / MALE |
| RIN | Adj Cov. (X5) | [2,...,7.9] |

In this model, the response variable was defined as follows:

- Response = YES: patients with partial response as their best response patients (n=6)
- Response= NO: patients with either progressive disease (PD) or stable disease (SD) as their best response and also patients with no tumour assessment available (n=96)

For each probe-set i, the aim of the statistical test was to reject the hypothesis that the mean expression levels in patients with response to treatment and patients without response to treatment are equal, taking into account the other adjustment covariates listed in table 2. Formally, the null hypothesis of equality was tested against a two sided alternative. Formally, the null hypothesis of equality was tested against a two sided alternative. Under the null hypothesis, the distribution of the t-statistic for this test follows a Student *t* distribution with 95 degrees of freedom. The corresponding p-values are reported in table 3.

The choice of linear model was motivated by two reasons. Firstly, linear modeling is a versatile, well-characterized and robust approach that allows for adjustment of confounding variables when estimating the effect of the variable of interest. Secondly, given the sample size of 102, and the normalization and scaling of the data set, the normal distribution assumption was reasonable and justified.

The issue of multiple testing was dealt with by using a False Discovery Rate (FDR) (Benjamini et al., Journal of the Royal Statistical Society Series B-Methodological, 1995. 57(1): p. 289-300) criterion for identifying the list of differentially expressed genes. Probe-sets with an FDR below the 0.3 threshold are declared significant. The 0.3 cut-off was chosen as a reasonable compromise between a rigorous correction for multiple testing with a stringent control of the risk of false positive and the risk of missing truly differential markers. The list of markers is reported in Table 3.

Table 3: Markers based on comparing "Responders" to "Non Responders".

Responders were defined as patients with Best Response equal to "Partial Response" (PR). Non Responders were defined as patients having "Stable Disease" (SD), "Progressive Disease" (PD) or no assessment available. Patients with no tumour assessment were included in the "Non Responder" group because in the majority of cases, assessment was missing because of early withdrawal due to disease progression or death.

Column 1 is the Affymetrix identifier for the probe-set. Column 2 is the GenBank accession number of the corresponding gene sequence. Column 3 is the corresponding

official gene name. Column 4 is the corresponding adjusted mean fold change in expression level between "responder" and "non responder". Column 5 is the p-value for the test of difference in expression level between "responders" and "non responders". Column 6 is the 95% confidence interval for the adjusted mean fold change in expression level.

| Affymetrix Probe Set ID | GenBank | Gene | Adjusted Mean Fold Change | P-value | CI 95% |
|---|---|---|---|---|---|
| 201816_s_at | NM_001483 | GBAS | 1.9 | 1.70E-04 | 1.4 , 2.7 |
| 205216_s_at | NM_000042 | APOH | 1.9 | 5.10E-05 | 1.4 , 2.6 |
| 205607_s_at | NM_020423<br>NM_181093 | SCYL3 | 1.5 | 4.90E-05 | 1.3 , 1.8 |
| 209805_at | NM_000535<br>NR_002217<br>NR_003085<br>XM_001126008<br>XR_017703 | PMS2CL | 1.3 | 1.60E-04 | 1.2 , 1.5 |
| 214203_s_at | NM_016335 | PRODH | 2.2 | 4.20E-05 | 1.5 , 3.0 |
| 215470_at | XM_001130621<br>XM_001130639<br>XM_001130651<br>XM_001130662<br>XM_001130670<br>XM_001130682 | DKFZP686M0199<br>SERF1A | 1.4 | 1.00E-05 | 1.2 – 1.6 |
| 216173_at | AK025360<br>NM_017920 | URG4 | 1.2 | 1.50E-04 | 1.1 , 1.3 |
| 220622_at | NM_024727<br>XM_001133921<br>XM_001133922<br>XM_001133923 | LRRC31 | 1.5 | 1.50E-05 | 1.3 , 1.8 |

For each probe-set, the assumption of homogeneity of variance was evaluated using Fligner-Killeen tests based on the model residuals. The analysis consisted of three steps:

Test all categorical variables for equality of residual variance between their levels

Note the variable V with the least p-value

If the least p-value is less than 0.001, re-fit the model allowing the different level of variables V to have a different variance.

Further statistical analysis

For the candidate markers GBAS, SCYL3 and SERF1A the following additional analyses were performed in a validated environment by an independent statisticians :

- Univariate Cox Regression for PFS (Progression free survival) from Primary Affymetrix Analysis,
- Univariate Logistic Regression for Response from Primary Affymetrix Analysis, and

The results of these analysis are presented below. They are consistent with the results of the primary analysis and confirm the choice of the selected marker.

Results: Univariate Cox Regression for PFS (Progression free survival) from Primary Affymetrix Analysis:

| **Gene** | **No. of patients** | **Hazard ratio** | **95 % CI for Hazard ratio** | **p-Value** |
|---|---|---|---|---|
| GBAS | 102 | 0.67 | 0.47; 0.95 | 0.0258 |
| SCYL3 | 102 | 0.36 | 0.19;0.68 | 0.0016 |
| SERF1A | 102 | 0.32 | 0.12;0.83 | 0.0191 |

Results: Univariate Cox Regression for Response from Primary Affymetrix Analysis:

| **Gene** | **No. of patients** | **Odds ratio** | **95 % CI for Odds ratio** | **p-Value** |
|---|---|---|---|---|
| GBAS | 102 | 15.02 | 2.68; 84.23 | 0.0021 |
| SCYL3 | 102 | >100 | 7.03;>1000 | 0.0011 |
| SERF1A | 102 | 56.04 | 4.79;656.22 | 0.0013 |

**Response to erlotinib treatment**

A total of 264 patients from 12 countries and 26 centres were enrolled in the study. 26% had Stage IIIB and 24% Stage IV NSCLC. 13.6% (n=36) of patients achieved an objective response while 31.4% (n=83) had clinical benefit (defined as having either an objective response or stable disease for 12 weeks or more). Median overall survival was 7.6 (CI 7-9) months and median progression-free survival was 11.3 (CI 8-12) weeks. Full details about the clinical data are shown in Table 1.

Fresh frozen bronchoscopic biopsies were collected from all subjects, but either not all samples had sufficient tumour content prior to microdissection (LCM) or did not have

sufficient RNA yield after LCM to proceed to microarray analysis, so that tumour material was only available for 125 patients; 122 of these had evaluable RNA. Another set of 20 samples did not pass our quality control assessment of the microarray data. Of the 102 microarray data sets that were suitable for statistical analysis, the clinical characteristics are shown in Table 1. While 36 patients in the overall study achieved an objective response, only 6 of these had microarray data; similarly for those achieving clinical benefit the number of subjects with microarray data was only 21 as compared to 83 in the full data set. 6 were judged to be partial responders (PR), 31 had SD and 49 had PD; of the 6 patients with a PR, 5 had adenocarcinoma and one had squamous cell carcinoma. There were no patients achieving a CR in the data set.

**Identification of genes associated with response to erlotinib**

Responders were defined as patients whose best response was partial response, while non-responders were defined as patients having either stable disease, progressive disease or for whom no assessment was made (in most cases as a result of early withdrawal due to disease progression or death). Thus in this model 6 "responders" were compared to 96 "non responders".

A linear model was fitted independently to each of the 16353 remaining probe-sets used in the analysis after removal of those probe-sets that were not present in any sample from the total 22283 on the HG-U133A microarray. A p-value was calculated for the difference in expression between response and non-response for each probe-set. A false discovery rate (FDR) of 0.3 was applied to correct for multiple testing. The list of 8 markers identified from this analysis is shown in Table 3.

**Discussion**

Targeting the Epidermal Growth Factor Receptor (EGFR) as a means of cancer therapy was proposed based on its ubiquitous aberrant expression in several epithelial cancers. EGFR is implicated in the pathogenesis and progression of many tumours including 40-80% of NSCLC tumours, as a result of activating mutations in the tyrosine kinase domain and / or its amplification. Upon activation, the receptor undergoes dimerization, resulting in phosphorylation of downstream targets with roles in cellular proliferation, metastasis, inhibition of apoptosis and neoangiogenesis.

Two major classes of EGFR inhibitors have been developed, monoclonal antibodies targeting the extracellular domain of the receptor, and small molecule tyrosine kinase inhibitors targeting the catalytic domain of the receptor. The latter include erlotinib which competes with ATP for the intracellular binding site.

It has emerged in recent years that several factors play a role in sensitivity to erlotinib including female gender, non-smoker status, Asian origin and adenocarcinoma histology; given that enhanced response rates are evident in such clinical subsets of patients, extensive efforts are ongoing to elucidate predictive molecular markers for patient stratification. Mutations in the *EGFR*, amplification of the *EGFR* gene locus and overexpression of EGFR on the protein level, have all been associated with response to varying degrees, though these are not the only molecular determinants of response.

By analyzing tissue samples with high-density oligonucleotide microarray technology, and applying statistical modeling to the data, we have been able to identify a set of eight genes whose expression levels are predictive of response to erlotinib (comparison of PR versus PD plus SD) (Table 3). Transcripts that are chromosomally located in the same region as the EGFR, including GBAS (1.9 fold upregulated; p = 0.00017) show a strong trend toward upregulation in the responders (comparison PR versus PD+SD). Such changes are suggestive of the presence of a chromosomal amplification around the EGFR gene locus of 7p11.2, which may be indicative of a good response to erlotinib. Amplification is a well-known mechanism exploited by tumour cells to increase the expression of a protein, activity of which promotes cell proliferation.

Glioblastoma amplified sequence or GBAS (located at 7p11.2) is a candidate marker that was found to be upregulated in PR as compared to PD+SD in our analyses (1.9 fold upregulated; p = 0.00017). Previous work has found GBAS to be co-amplified with EGFR in two out of 12 glioblastomas as well as in 2 of 3 cell lines; the gene was not amplified in glioblastoma tissues lacking EGFR amplification, suggesting co-amplification of a larger region. Additional work from the same group suggests that EGFR amplicons can exceed 1Mb in length and may be substantially longer reaching up to 5Mb. Thus this would support the notion of coamplification of a larger stretch of the cytoband around 7p11.2.

Apolipoprotein H (APOH) which was expressed 1.9 fold higher in PR as compared to PD (p=0.000051) has been linked to aggressive non-Hodgkin's lymphoma where antibodies to this protein and other phospholipids may be a prognostic marker.

SCY1-like 3 (SCYL3) codes for a ubiquitously-expressed protein known to interact with ezrin, an adhesion receptor molecule involved in regulating cell shape, adhesion, motility and responses to the extracellular environment (Sullivan et al, 2003).

Table 4: List of marker genes of the present invention

Column 1 is the GenBank accession number of the human gene sequence; Column 2 is the corresponding official gene name and Column 3 is the Sequence Identification number of the human nucleotide sequence as used in the present application. For certain genes table 4 contains more than one sequence identification number since several variants of the gene are registered in the GeneBank.

| **GenBank Accession number** | **Gene** | **Sequence identification number** |
|---|---|---|
| NM_001483 | GBAS | Seq. Id. No. 1 |
| NM_000042 | APOH | Seq. Id. No. 2 |
| NM_020423<br>NM_181093 | SCYL3 | Seq. Id. No. 3<br>Seq. Id. No. 4 |
| NM_000535<br>NR_002217<br>NR_003085<br>XM_001126008<br>XR_017703 | PMS2CL | Seq. Id. No. 5<br>Seq. Id. No. 6<br>Seq. Id. No. 7<br>Seq. Id. No. 8<br>Seq. Id. No. 9 |
| NM_016335 | PRODH | Seq. Id. No. 10 |
| XM_001130621<br>XM_001130639<br>XM_001130651<br>XM_001130662<br>XM_001130670<br>XM_001130682 | DKFZP686M0199<br>SERF1A | Seq. Id. No. 11<br>Seq. Id. No. 12<br>Seq. Id. No. 13<br>Seq. Id. No. 14<br>Seq. Id. No. 15<br>Seq. Id. No. 16 |
| AK025360<br>NM_017920 | URG4 | Seq. Id. No. 17<br>Seq. Id. No. 18 |
| NM_024727 | LRRC31 | Seq. Id. No. 19 |

## Claims

1. An *in vitro* method of predicting the response of a cancer patient to treatment with an EGFR inhibitor comprising: determining the expression level of at least one gene selected from the group consisting of GBAS, APOH, SCYL3, PMS2CL, PRODH, SERF1A, URG4A and LRRC31 in a tumour sample of a patient and comparing the expression level of the at least one gene to a value representative of an expression level of the at least one gene in a non responding patient population, wherein a higher expression level of the at least one gene in the tumour sample of the patient is indicative for a patient who will respond to the treatment.

2. The method of claim 1, wherein the expression level is determined by microarray technology.

3. The method of claim 1 or 2, wherein the expression level of at least two genes is determined.

4. The method of claims 1 to 3, wherein the expression level of at least three genes is determined.

5. The method of claims 1 to 4, wherein the EGFR inhibitor is erlotinib.

6. The method of claims 1 to 5, wherein the cancer is NSCLC.

7. Use of a gene selected from the group consisting of GBAS, APOH, SCYL3, PMS2CL, PRODH, SERF1A, URG4A and LRR 31 for predicting the response of a cancer patient to EGFR inhibitor treatment.

8. The use of claim 7, wherein the cancer is NSCLC.

9. The use of claim 7 or 8, wherein the EGFR inhibitor is erlotinib.

10. A method of treating a cancer patient identified by a method of claims 1 to 7 comprising administering an EGFR inhibitor to the patient.

11. The method of claim 10, wherein the EGFR inhibitor is erlotinib.

12. The method of claim 10 or 11, wherein the cancer is NSCLC.

Fig. 1

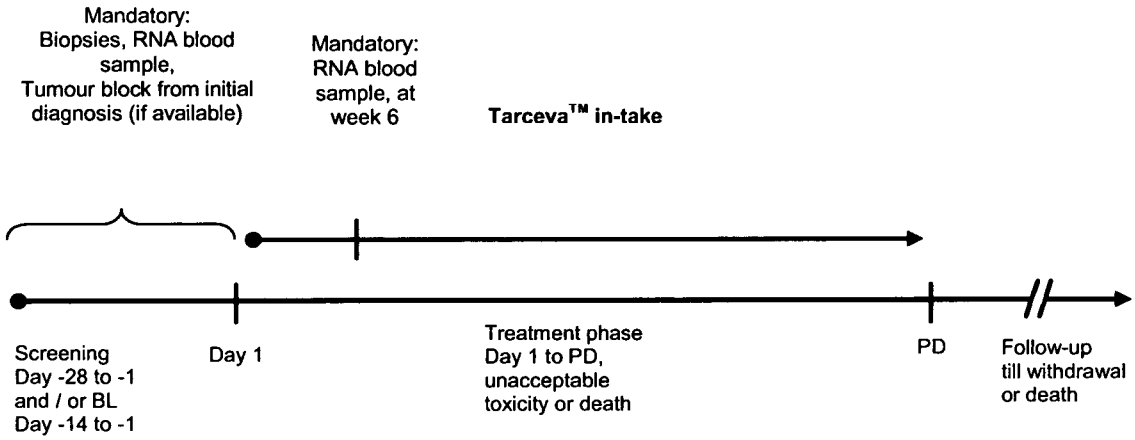
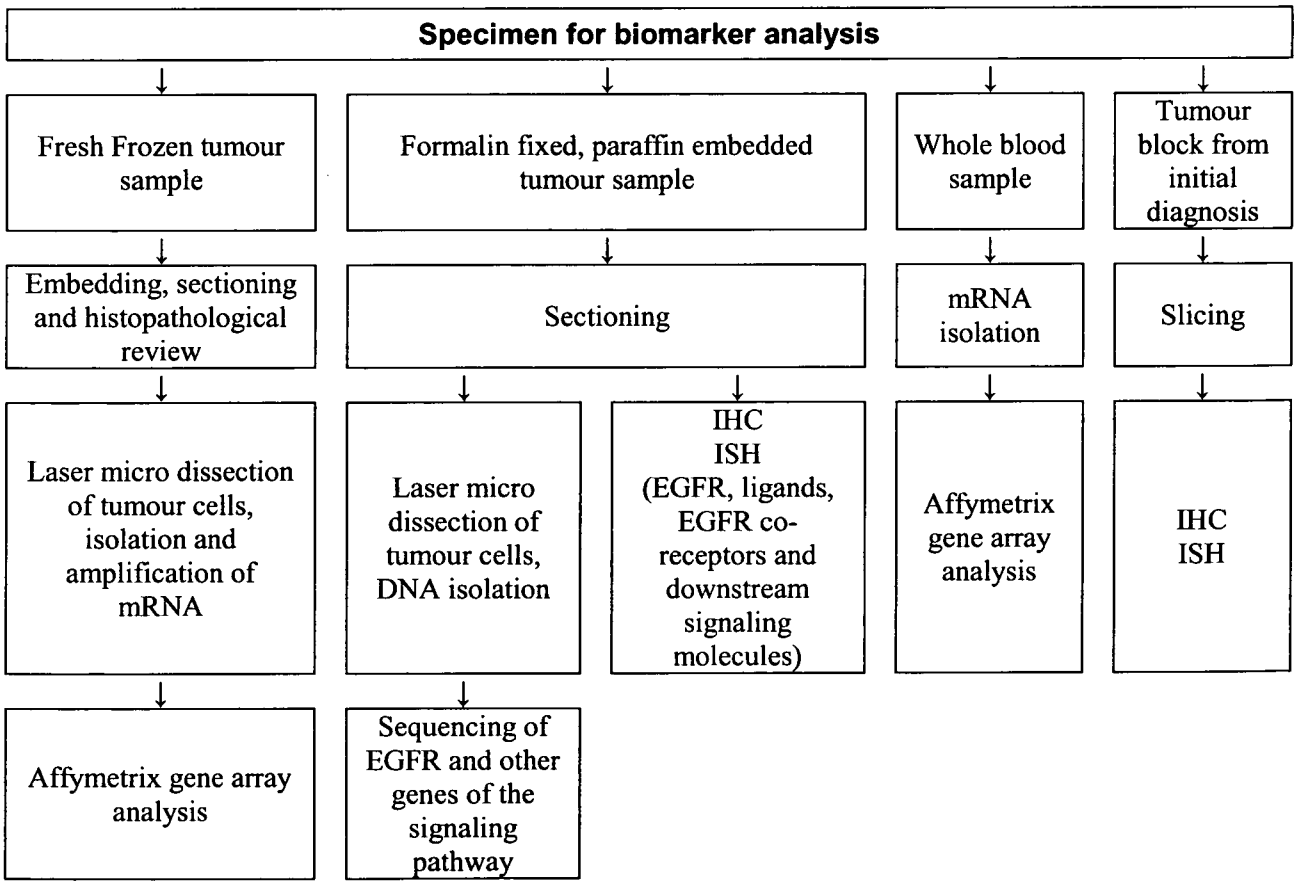

Mandatory:
Biopsies, RNA blood sample,
Tumour block from initial diagnosis (if available)
Mandatory:
RNA blood sample, at week 6
Tarceva™ in-take
Screening
Day -28 to -1
and / or BL
Day -14 to -1
Day 1
Treatment phase
Day 1 to PD,
unacceptable
toxicity or death
PD
Follow-up
till withdrawal
or death

Fig. 2


Specimen for biomarker analysis
Fresh Frozen tumour sample
Formalin fixed, paraffin embedded tumour sample
Whole blood sample
Tumour block from initial diagnosis
Embedding, sectioning and histopathological review
Sectioning
mRNA isolation
Slicing
Laser micro dissection of tumour cells, isolation and amplification of mRNA
Laser micro dissection of tumour cells, DNA isolation
IHC
ISH
(EGFR, ligands, EGFR co-receptors and downstream signaling molecules)
Affymetrix gene array analysis
IHC
ISH
Affymetrix gene array analysis
Sequencing of EGFR and other genes of the signaling pathway

# INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/006512

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, MEDLINE, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 2006/119980 A (HOFFMANN LA ROCHE [CH]; BRENNSCHEIDT SABINE [DE]; HELLER ASTRID [DE];) 16 November 2006 (2006-11-16) page 22 page 16 | 1-12 |
| X | WO 2006/101925 A (OSI PHARM INC [US]; GENENTECH INC [US]; HALEY JOHN D [US]; GRIFFIN GRA) 28 September 2006 (2006-09-28) page 21; claims 1,2 | 1-12 |
| X | WO 2005/049829 A (ASTRAZENECA UK LTD [GB]; UNIV TOKYO [JP]; TSURUO TAKASHI [JP]; NAKAMUR) 2 June 2005 (2005-06-02) page 2 page 8; claim 13 | 1-12 |

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 5 November 2008 | 29/01/2009 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL – 2280 HV Rijswijk<br>Tel. (+31–70) 340–2040,<br>Fax: (+31–70) 340–3016 | Reuter, Uwe |

3

Form PCT/ISA/210 (second sheet) (April 2005)

# INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2008/006512

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 2004/071572 A (GENOMIC HEALTH INC [US]; CEDARS SINAI MEDICAL CENTER [US]; AGUS DAVID) 26 August 2004 (2004-08-26) pages 14-16; claims | 1-12 |
| X | MING-SOUND TSAO ET AL: "Erlotinib in Lung Cancer Molecular and Clinical Predictors of Outcome" NEW ENGLAND JOURNAL OF MEDICINE, MASSACHUSETTS MEDICAL SOCIETY, BOSTON, MA, US, vol. 353, no. 2, 14 July 2005 (2005-07-14), pages 133-145, XP007905868 ISSN: 1533-4406 abstract | 1-12 |
| X | WO 2007/067500 A (GENOMIC HEALTH INC [US]; BAKER JOFFRE B [US]; KIEFER MICHAEL C [US]) 14 June 2007 (2007-06-14) paragraph [0070]; claim 28 paragraph [0051] | 1-12 |

3

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2008/006512

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers allsearchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search reportcovers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1 - 12 (all partially)

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of GBAS, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

   ---

2. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of APOH, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

   ---

3. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of SCYL3, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

   ---

4. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of PMS2CL, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

   ---

5. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of PRODH, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

   ---

6. claims: 1-12 (all partially)

   An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of SEF1A, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

---

7. claims: 1-12 (all partially)

An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of URG4a, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

---

8. claims: 1-12 (all partially)

An in vitro method for predicting the response to the treatment with an EGFR inhibitor comprising determining the expression level of LRRC31, the use of said gene to predict response to EGFR inhibitor treatment and a method of treatment of a patient identified by said method.

---

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2008/006512

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 2006119980 A | 16-11-2006 | AR 053272 A1 | 25-04-2007 |
| | | AU 2006245962 A1 | 16-11-2006 |
| | | CA 2605151 A1 | 16-11-2006 |
| | | JP 2008538817 T | 06-11-2008 |
| | | KR 20070116641 A | 10-12-2007 |
| | | NO 20074389 B | 06-12-2007 |
| | | US 2007054330 A1 | 08-03-2007 |
| WO 2006101925 A | 28-09-2006 | CA 2601157 A1 | 28-09-2006 |
| | | EP 1861715 A2 | 05-12-2007 |
| | | JP 2008533490 T | 21-08-2008 |
| WO 2005049829 A | 02-06-2005 | AU 2004291709 A1 | 02-06-2005 |
| | | BR PI0410634 A | 13-06-2006 |
| | | CA 2527680 A1 | 02-06-2005 |
| | | EP 1633870 A1 | 15-03-2006 |
| | | JP 2006526420 T | 24-11-2006 |
| | | KR 20060002012 A | 06-01-2006 |
| | | MX PA05012939 A | 17-05-2006 |
| | | US 2006252056 A1 | 09-11-2006 |
| WO 2004071572 A | 26-08-2004 | AU 2004211955 A1 | 26-08-2004 |
| | | CA 2515096 A1 | 26-08-2004 |
| | | EP 1590487 A2 | 02-11-2005 |
| | | JP 2006521793 T | 28-09-2006 |
| WO 2007067500 A | 14-06-2007 | NONE | |